FLUENT AND HESITATION PAUSES AS A FUNCTION OF SYNTACTIC COMPLEXITY

By

KENNETH F. RUDER

A DISSERTATION PRESENTED TO THE GRADUATE COUNCIL OF
THE UNIVERSITY OF FLORIDA
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE
DEGREE OF DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA
1969

To Cee Cee, Christine
and Maribeth

ACKNOWLEDGMENTS

The author expresses deep gratitude to his advisor, chairman and friend, Dr. Paul Jensen, for his continued support and guidance, not only throughout the planning and execution of this investigation, but also throughout the author's academic career at the University of Florida. The author is further indebted to the members of the dissertation committee for their interest, constructive criticism and special contributions to the ideas expressed herein.

The author further wishes to acknowledge the invaluable support of the faculty, staff and students of the Communication Sciences Laboratory of the University of Florida.

The author also expresses his appreciation and gratitude to his wife and family who provided steady encouragement, inspiration and support.

Finally, the author expresses his appreciation to the National Science Foundation (Grant GS-2337) and the National Institute of Neurological Diseases and Stroke (Grant NB-06459) for their financial support in this investigation.

TABLE OF CONTENTS

Page

LIST OF TABLES.. vi

LIST OF FIGURES... viii

CHAPTER

I INTRODUCTION AND STATEMENT OF THE PROBLEM... 1
  Introduction.............................. 1
  Statement of the Problem.................. 7

II PROCEDURE................................. 9
  Subjects.................................. 10
  Stimulus Sentences........................ 11
  Preparation of Stimulus Sentences......... 16
  Experimental Apparatus.................... 18
  Perceptual Tasks.......................... 22
  Pause Duration Measurement................ 24
  Statistical Analysis...................... 27

III RESULTS AND DISCUSSION.................... 29
  Comparison of Durations for Pause Detection Threshold, Fluent Pause and Hesitation Pause.................................. 29
    Within-Phrase Pause Durations.......... 31
    Between-Phrase Pause Durations......... 37
  Effects of Syntactic Complexity on Pause Durations............................... 42
    Pause Detection as a Function of Syntactic Complexity.................. 46
    Fluent Pause Duration as a Function of Syntactic Complexity.................. 49
    Hesitation Pause as a Function of Syntactic Complexity.................. 51
  Subject Variability....................... 53

IV SUMMARY AND CONCLUSIONS.................... 61

Page

APPENDIX A INSTRUMENTATION AND PROCEDURE USED IN THE DURATION MATCHING TASK OF THE SUBJECT SCREENING TEST................. 67

APPENDIX B TREE DIAGRAM CONSTITUENT STRUCTURE ANALYSIS OF THE FIVE STIMULUS SENTENCES............................ 70

APPENDIX C INSTRUCTION TO SUBJECTS AND SENTENCES USED IN THE SYNTACTIC COMPLEXITY RANKING TASK.......................... 76

APPENDIX D INSTRUCTIONS AND RESPONSE FORM FOR THE SPEAKER ATTRIBUTE RANKING.............. 79

APPENDIX E GENERATION OF THE TEST TAPES........... 81

APPENDIX F PAUSE DURATION DATA.................... 85

APPENDIX G MEAN PAUSE DURATION AND STANDARD DEVIATION AT EACH LEVEL OF SYNTACTIC COMPLEXITY............................ 89

APPENDIX H MEAN PAUSE DURATION AND STANDARD DEVIATION FOR EACH SYNTACTIC COMPLEXITY LEVEL AND TASK........................ 93

BIBLIOGRAPHY... 95

BIOGRAPHICAL SKETCH.................................. 97

LIST OF TABLES

Table Page

1. Summary of performances (in msec.) on the pure tone duration matching task employed in the subject screening test................ 12

2. List of stimulus sentences. The parenthesized numbers are the structural complexity index (Miller and Chomsky, 1963) values for the syntactic relations between "lost" and "contact."................................... 13

3. Subjects' rank orderings of the complexity of the syntactic relations between the words "lost" and "contact" in the five stimulus sentences (listed in Table 2)................ 15

4. Summary of treatments by subjects analysis of variance to determine the effects of task, syntactic complexity and replications on pause duration........................... 32

5. Simple main effects for factor A (tasks) at each level of factor B (syntactic complexity) 33

6. Simple main effects for factor B (syntactic complexity) at each level of factor A (tasks)....................................... 34

7. Comparison of between-phrase pause durations at sentence three for pause detection threshold, fluent pause and hesitation pause. The marginal values are the ordered pause duration means. The values in the matrix represent the differences between each marginal pair........................... 38

Table Page

8. Comparison of the between-phrase pause durations at sentence four for pause detection threshold, fluent pause and hesitation pause. The marginal values are the ordered means. The values in the matrix represent the differences between each marginal pair.............................. 39

9. Comparison of the between-phrase pause durations at sentence five for pause detection threshold, fluent pause and hesitation pause. The marginal values are the ordered means. The values in the matrix represent the differences between each marginal pair.............................. 40

10. Comparison of hesitation pause duration means for each level of syntactic complexity. The marginal values contain the ordered mean durations for each level of complexity. The values in the matrix represent the differences between each marginal pair....... 45

11. Summary of the analysis of variance used to test for between subject effects.......... 57

12. Summary of analysis of variance used to test for differences between high and low variable groups of subjects.................. 59

13. Pause duration data for replication one...... 86

14. Pause duration data for replication two...... 87

15. Pause duration data for replication three.... 88

16. Mean and standard deviation of pause detection thresholds.............................. 90

17. Mean and standard deviation of fluent pause durations.................................. 91

18. Mean and standard deviation of hesitation pause durations............................. 92

19. Mean and standard deviation of pause detection threshold, fluent pause and hesitation pause durations at each level of syntactic complexity..................... 94

LIST OF FIGURES

Figure		Page
1.	Mean rank of speakers on six attributes......	17
2.	Instrumentation for variable delay system to adjust pause durations....................	19
3.	Block diagram of instrumentation for experimental pause adjustment and measurement......	21
4.	Reference points used in the measurement of pause durations from the oscillographic photographs: line AB, the top of the Channel A baseline; line EF, a perpendicular erected through AB at point X (point where the speech signal drops to the baseline AB); line CD, the bottom of the Channel B baseline; line GH, a perpendicular erected at point Y (point at which the Channel B integrated signal drops below the baseline); line IJ, the distance in centimeters from the end of Channel A signal to the beginning of Channel B signal..	25
5.	Comparison of mean durations of pause detection threshold, fluent pause and hesitation pause.	30
6.	Comparison of sentence one and two mean pause durations for pause detection threshold, fluent pause and hesitation pause............	36
7.	Mean pause duration for each of five levels of increasing syntactic complexity...........	43
8.	Mean pause duration of the five levels of increasing syntactic complexity for each task	44
9.	A comparison of fluent pause durations with pause detection thresholds obtained from measures in which the plosive burst terminating the word "lost" was included as part of the speech signal and pause detection thresholds obtained from measures in which this plosive burst was ignored...........................	48

Figure Page

10. Overall mean pause duration for each of three replications.................................. 54

11. Mean pause duration for replications and tasks at each level of syntactic complexity........ 55

12. Block diagram of instrumentation used in the duration matching task of the subject screening test.. 68

13. Tree diagram analysis of stimulus sentence one. The node-to-terminal-node ratio specifying the complexity of the relation between "lost" and "contact" is 3/2 (1.5)...................... 71

14. Tree diagram analysis of stimulus sentence two. The node-to-terminal-node ratio specifying the complexity of the relation between "lost" and "contact" is 7/4 (1.75)..................... 72

15. Tree diagram analysis of stimulus sentence three. The node-to-terminal-node ratio specifying the complexity of the relation between "lost" and "contact" is 13/7 (1.85).. 73

16. Tree diagram analysis of stimulus sentence four. The node-to-terminal-node ratio specifying the complexity of the relation between "lost" and "contact" is 15/8 (1.88).. 74

17. Tree diagram analysis of stimulus sentence five. The node-to-terminal-node ratio specifying the complexity of the relation between "lost" and "contact" is 27/14 (1.93).. 75

18. Block diagram of instrumentation used in the generation of the two-track test tape......... 82

CHAPTER 1

INTRODUCTION AND STATEMENT OF THE PROBLEM

Introduction

One of the most prevalent, yet most overlooked, aspects in the study of speech behavior concerns those events which interrupt the continuous flow of normal speech. These events have traditionally been referred to in the literature as speech pauses. Although such events are characteristic of all speakers, there are actually very few data available concerning their nature and function. This is not to say that pausal behavior in speech has not been studied. There has been a considerable number of investigations (Boomer and Dittman, 1962; Goldman-Eisler, 1961a and 1961b; Hargreaves and Starkweather, 1959; Lynch, 1934; and Agnello, 1963) which have attempted to identify the physical and perceptual parameters of speech pauses, but these have met with only limited success. In fact, these studies have yielded results often incompatible and inconsistent with each other, primarily because they were based on observation of pauses as they occur spontaneously rather than on rigorously controlled experimentation. As a result, the nature and function of speech pauses remain generally obscure.

Actually, it is not even clear as to what constitutes a pause in speech. A speech pause has been defined, on one hand, as the time interval between the cessation of a sound and the beginning of the succeeding sound (Agnello, 1963; Black, et al., 1966; Aaronson and Markowitz, 1967; Boomer and Dittman, 1962; Judson and Weaver, 1965; Lynch, 1934). On the other hand, there are those who focus on the allophonic changes and suprasegmental variations associated with linguistic juncture as basic parameters of the speech pause (Trager and Smith, 1951; Jones, 1962; Pike, 1945; Bloomfield, 1933). The neutral position is exemplified by a definition of speech pause as a complex interaction between allophonic changes and silent intervals (Mattingly, 1966).

In any case, it does seem clear that the silent interval between two segmental phonemes is a primary parameter cueing the perception of a pause. For example, Agnello (1963) reported that those events perceived as pauses by speaker-listeners were marked by silent intervals of longer durations than those silent intervals not perceived as pauses. Goldman-Eisler (1961b) noted that events perceived as hesitation pauses were marked likewise by silent intervals of relatively long duration. These two studies, as well as those by Lounsbury (1954) and Hargreaves and Starkweather (1959), can thus be interpreted as demonstrating that there is a high correlation between what is perceived as a pause and silent intervals of relatively long duration.

Nevertheless, there is a lack of direct evidence which explicates the relationship between the physical event (silent interval) and its perception in speech. Several types of information, based upon systematic and controlled experimentation, are needed. A basic type concerns the variability in the perception of such intervals. The literature is completely lacking any data in this regard. One might infer, however, that since there is considerable variability in the use and distribution of silent intervals in speech (Goldman-Eisler, 1961a and 1961b), there is a similar variability in their perception. A speculation of this nature (namely, that the psychological value of a pause of a given duration is highly variable) has been advanced (Judson and Weaver, 1965), but to date it remains to be empirically verified.

Information is needed also relative to whether silent intervals are perceived as different types of pauses. The literature indicates that this question should be answered affirmatively. There seems to be no question but that individuals are able to perceive two types of pauses in speech (Boomer and Dittman, 1962), those which are identified as hesitations and those which are not. For lack of better terms, the former will be referred to as hesitation pauses and the latter as fluent pauses.[1]

Implicit in this perceptual dichotomy of speech pauses is the assumption that pause duration is a primary cue responsible for the differential perception. As previously stated, it has often been asserted (Maclay and Osgood, 1959; Lieberman, 1967; Goldman-Eisler, 1961b; Boomer and Dittman, 1962) that hesitation pauses are silent intervals of unusual length (duration). Unfortunately, also as previously stated, the best evidence that is offered in support of this assertion is that pauses perceived as hesitations tend to be associated with silent intervals of durations greater than those associated with fluent pauses. However, it should be noted that in such studies, the phonetic (and to some extent, the syntactic and semantic) environment differed for each pause comparison made. These varying environments may have served as perceptual cues of hesitation as well as, or instead of, the duration of the silent interval. Thus, it is apparent that the concept of the silent interval as a primary cue signaling the perception of hesitation or fluent pauses needs further examination with environmental variables being controlled.

---

[1] It is conventional in the literature to refer to silent intervals, other than hesitations, as pauses. The opposition of fluent and hesitation pauses is therein implied, but here is stated directly for purposes of clarity.

Even then, a further consideration needs to be made. In comparing the studies which have claimed that the perceptual distinction between fluent and hesitation pauses is duration, one finds that there is considerable disagreement concerning the perceptual value of a pause of a given duration. Lounsbury (1954) reports that pauses of durations greater than 100 msec. are perceived as hesitations, while Goldman-Eisler (1961b) sets 250 msec. as the lower bound of hesitation pauses. Hargreaves and Starkweather (1959), on the other hand, state that perceptual identifications of hesitation pauses were consistent only for silent intervals of durations greater than one second. Boomer and Dittman (1962) dispute this finding and indicate that their subjects were able to discriminate hesitation pauses reliably at durations down to 200 msec. Thus, the durational specification of silent intervals perceived as being hesitation pauses needs clarification.

This problem is even more acute in the study of fluent pauses. There have been, to date, no direct attempts at specifying the durational bounds for fluent pauses, although one may infer these from the hesitation pause thresholds. Following this line of reasoning, it might be assumed that the durational domain of the fluent pause ranges from the point at which a silent interval is just perceived, to that point where the silent interval is perceived as being a hesitation. However, in view of the contradictory evidence previously discussed with regard to this latter point, it is evident that any measures of fluent pauses derived from hesitation pause thresholds will be as inconsistent as the data on which they were based.

Furthermore, there is no information available concerning the former point, the threshold of pause detection in speech, even though this threshold has been implicitly designated as the lower bound of fluent pauses (Agnello, 1963). It is not evident, however, why this should necessarily be the case since one could just as well assume that the pause detection threshold is essentially identical in duration to those silent intervals perceived as fluent pauses and that anything in excess of this duration is perceived as a hesitation. Comparison of Agnello's (1963) minimum duration of perceived fluent pauses of 190 msec. to Boomer and Dittman's (1962) minimum duration of perceived hesitation pauses of 200 msec. could lead one to make such an assumption. In any case, the study of the durational specifications of those silent intervals perceived as fluent pauses cannot profitably proceed until more reliable data about the threshold of hesitation pauses and the threshold of pause detection are obtained.

In view of the widespread disagreement and inconsistent data relating to the durational specification of fluent and hesitation pauses, it would seem reasonable to consider the viewpoint that duration of the silent interval alone is not an adequate perceptual cue in differentiating fluent and hesitation pauses. Nevertheless, there have been no systematic attempts to locate and study other cues which might contribute to the differential perception of speech pauses. One highly likely cue, as mentioned earlier, which might interact with duration of the silent interval to provide this perceptual differentiation, is the syntactical environment which bounds the silent interval. Syntactical environment, in this sense, refers to the relationships which exist between the various constituent constructions of the sentence. Such relationships, for example, may be expressed in terms of the conventional immediate constituent analysis used in linguistic descriptions.

There is nothing novel about postulating a relationship between syntactic structure and duration of the silent interval as providing perceptual cues for speech pauses. Such a relationship is implicit in those reports which seek to relate the function of fluent pauses to signaling the syntactic structure of a sentence (Carrell and Tiffany, 1960; Jones, 1962; Lieberman, 1967). However, this relationship is at present nothing more than a broad sweeping generalization which has yet to be verified. If true, it would require that the duration of the silent interval be directly related to the complexity of the syntactic structure which it signals since duration is the only parameter which can be manipulated to signal this structure. That is to say, such a viewpoint would have to claim that the absolute duration of the silent interval is not the primary perceptual cue for speech pauses. Rather, it would be the duration of this interval, relative to the durations of the other silent intervals in the sentence, which would be perceptually significant.

It is likely then, if fluent pauses vary in duration relative to the complexity of the syntactic structure, that the duration of hesitation pauses might likewise vary as a function of syntactic structure. Such a conception might well help explain the discrepancies found in previous studies of hesitation pause duration by claiming that the disparate measures resulted from observing pauses in different syntactical environments. Furthermore, if duration does indeed vary with syntactic complexity to provide the perceptual cues for fluent and hesitation pauses, then it might well be that these cues also interact to provide pause detection thresholds which vary as a function of syntactic complexity.

Statement of the Problem

Previous studies have indicated that, perceptually, there are two types of pauses in speech, fluent pauses and hesitation pauses. These studies, moreover, have focused on the duration of the silent interval as being the only significant factor contributing to the differential perception of such pauses. As pointed out above, however, there is evidence which indicates that duration is not an independent parameter contributing to the perceptual differentiation of speech pauses. At least one other variable, syntactic structures bounding the silent interval, has been postulated as playing a role in their perception (Boomer and Dittman, 1962). It seems reasonable then to postulate a description of perceptual processing of speech pauses which is dependent not only on duration, but also on the syntactical environment bounding the silent interval. There is, however, no evidence of an empirical nature which directly substantiates such a claim. Finally, although a number of investigators have attempted to determine the perceptual nature of speech pauses in relation to silent intervals, as such, there are few or no direct data available concerning the variability of the perception of speech pauses.

It is in regard to these considerations, therefore, that the present study has been formulated to provide information concerning the following questions:

(1) Does duration of the silent interval differ for pause detection thresholds, fluent pauses and hesitation pauses within identical environments?

(2) Does duration of the silent interval for pause detection threshold, fluent pause and hesitation pause vary as a function of syntactic complexity?

(3) How variable are subjects' perceptions of the duration of the silent intervals for pause detection threshold, fluent pause and hesitation pause?

CHAPTER II

PROCEDURE

The purpose of the study was to investigate the effects of syntactic environment and duration of the silent interval on the perception of pauses in speech. To accomplish this purpose, the following general approach was developed. The stimuli consisted of five sentences within which the words "lost" and "contact" were manipulated so that their syntactic relation to one another varied in complexity from sentence to sentence. Each stimulus sentence was then recorded so that the portion of a sentence following the word "lost" and preceding the word "contact" was on Track A of a two-track tape, while the remainder of the sentence was recorded on Track B. Twelve subjects, working individually, adjusted the duration of the silent interval between "lost" and "contact" within each sentence in order to locate (1) the point at which the silent interval was just detected as a pause (pause detection threshold), (2) the point which was considered optimal for the fluent presentation of the sentence (fluent pause) and (3) the point at which the pause was just perceived as being a hesitation (hesitation pause). Details concerning the specific procedures are presented in the following paragraphs.

Subjects

Twelve volunteer undergraduate college students were selected as subjects for the investigation. They were chosen from an initial pool of fifteen individuals on the basis of the following criteria: (1) all were native speakers of English; (2) they were free of any noticeable speech and/or hearing defect; (3) they were naive regarding the literature on hesitation phenomena or pausal behavior in speech (that is, they had not read about or participated in experiments or studies concerning pausal behavior of any kind); and (4) they were able to make temporal comparisons within a prescribed degree of accuracy.

The need for criteria one and two is self-evident; criterion three was imposed in order to avoid possible bias resulting from such experience. The assessment as to whether an individual satisfied the first three subject-selection criteria was made informally during a short screening test. Only one individual, who reported a long-standing monaural hearing loss, failed this portion of the test.

Criterion four was adopted in order to reduce the variability present in the criterion measures of pause durations due to a particular subject's general difficulty in making temporal judgments. It was applied as follows. The individual was asked to adjust the duration of a 250 Hz pure tone, taking as much time as necessary, to match that of a standard tone (also 250 Hz) of 400 msec. duration.[2] Three such duration matches were required of each individual.

---

[2] A description of the instrumentation and procedure used in this part of the study is presented in Appendix A.

Satisfactory performance on this part of the screening test required that the mean of the three duration matches be within plus or minus 72 msec.[3] of the standard tone duration. Two of the remaining fourteen individuals failed to meet this criterion and were rejected as subjects for the study. Table 1 presents a summary of the duration matches made by each of the resultant twelve individuals who served as subjects.

Stimulus Sentences

The stimulus sentences are listed in Table 2. In these sentences, the words "lost" and "contact" were manipulated so that they assumed different syntactic relations to one another. These syntactic relations were specified by means of tree diagrams (see Appendix B) which defined the constituent structure of each stimulus sentence.

Since one of the purposes of the study was to study the effect of syntactic environment on the perception of speech pauses, it was necessary to quantify the complexity of the syntactic relations between "lost" and "contact" so that these relations which constitute the syntactic environments could be rank ordered. This quantification was accomplished through use of a structural complexity index (Miller and Chomsky, 1963).

To apply this metric, one locates the lowest node of the tree diagram which encompasses both words of interest, "lost" and "contact,"

---

[3] Small and Campbell (1962) report that the absolute difference limen for a 250 Hz tone with a reference duration of 400 msec. is 72 msec. For purposes of this screening criterion, it was felt that the subjects' accuracy should at least approximate this difference limen value.

Table 1. Summary of performances (in msec.) on the pure tone duration matching task employed in the subject screening test.

Subject	Duration Match I	Duration Match II	Duration Match III	Mean	S.D.
1	302	470	402	391	69
2	430	340	414	395	39
3	383	478	395	419	42
4	395	411	326	377	37
5	412	391	407	403	0.06
6	320	510	391	407	78
7	422	408	451	427	18
8	355	385	359	366	13
9	520	441	409	457	47
10	458	395	341	398	48
11	288	399	461	383	72
12	471	434	388	431	34

Table 2. List of stimulus sentences. The parenthesized numbers are the structural complexity index (Miller and Chomsky, 1963) values for the syntactic relations between "lost" and "contact."

---

1. He lost (1.5) contact with reality.
2. The lost (1.75) contact lens case was finally found.
3. The team that lost (1.85) contacted the authorities.
4. If you are lost (1.88) contact the nearest policeman.
5. Even though the battle was lost (1.93) contact had been established with the rear guard.

and counts the total number of nodes contained in this branch of the tree structure. This number is then divided by the number of terminal nodes contained in that branch. For instance, in the sentence, "He lost contact with reality," the lowest node which encompasses the words "lost" and "contact" is the node labeled 1 (see Appendix B). The total number of nodes in this branch is three, and the number of terminal nodes is two. The node-to-terminal-node ratio of 3/2 (1.5) is thus the structural complexity index for that syntactic relation between the words "lost" and "contact." The structural complexity index values for the different syntactic relations between "lost" and "contact" are given in the parentheses in Table 2.

Since the constituent structure analysis used in arriving at the tree diagram is an arbitrary procedure in many instances, it was decided to obtain another measure of the complexity of the syntactic relations between the words "lost" and "contact" which might be used as an external test of the validity of the structural complexity index as used in this study. To this end, eighteen native speakers of English with no formal training in linguistics were asked to rank order the complexity of the syntactic relations between the words "lost" and "contact" and "past (passed)" and "quick." The sentences containing the words "lost" and "contact" were the five stimulus sentences; those containing "past (passed)" and "quick" were five foil sentences. Appendix C contains the instructions to the subjects and the list of sentences used in this ranking procedure.

Subject rankings of complexity of the syntactic relations between the words "lost" and "contact" are presented in Table 3. A Spearman Rank Order Correlation Coefficient for the relation between the

Table 3. Subjects' rank orderings of the complexity of the syntactic relations between the words "lost" and "contact" in the five stimulus sentences (listed in Table 2).

	Sentence				
Subject	1	2	3	4	5
1	1	2	3	4	5
2	2	1	5	3	4
3	1	3	4	2	5
4	1	2	3	4.5	4.5
5	2	1	5	3	4
6	2	1	4	3	5
7	2	1	3	4	5
8	1	2	3	4	5
9	1	4	5	2	3
10	1	2	3	5	4
11	3	1	2	4.5	4.5
12	1	2	3	4	5
13	1	3	2	5	4
14	3	1	2	4.5	4.5
15	1	2	3	4	5
16	1	2	3	4	5
17	1	5	3.5	3.5	2
18	3	1	2	4	5
Mean Rank	1.56	2.0	3.25	3.83	4.36

structural complexity index rank and the eighteen subjects' rank orderings was computed and found to be .96. A Kendall Coefficient of Concordance of .58 (significant beyond the .01 level) also indicates that there was good intersubject agreement on the syntactic complexity ranking task.

Thus, the Spearman rho of .96 was taken as evidence that the structural complexity index as used with the five sentences in this study was an adequate and valid measure of syntactic complexity. The levels of complexity of the syntactic relations between the words "lost" and "contact" in the five stimulus sentences thus ranged from sentence one containing the least complex to sentence five containing the most complex relations.

Preparation of Stimulus Sentences

Four adult male speakers were recorded reading the five stimulus sentences from individual 3 x 5 cards. The four recordings of each sentence were presented together to fifteen students in an advanced speech pathology class for ranking of the speakers with respect to the following attributes: (1) quality, (2) fluency, (3) naturalness, (4) rate, (5) precision and (6) overall effectiveness. Appendix D contains the instructions and response form used in this task.

Figure 1 presents a graphic review of the results of this procedure. As can be seen, Speaker A was ranked considerably better on all attributes than any of the other three speakers. The five sentences as recorded by Speaker A, therefore, were selected for use in the experiment.

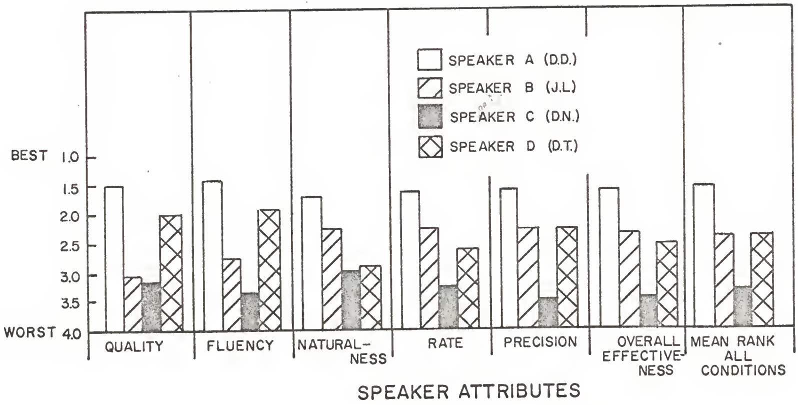

Figure 1. Mean rank of speakers on six attributes.

Speaker A's recordings of each of the five stimulus sentences were then dubbed through a two-channel tape recorder in such a manner that the words preceding the silent interval between "lost" and "contact" were recorded on Track A and the words following this interval were recorded on Track B. These two-track tapes were, in turn, redubbed repeatedly so that each of the five sentences was recorded on an individual reel of tape approximately 150 times. A five-second inter-stimulus interval separated each repetition. A detailed description of the procedure for the generation of these two-track tapes, one for each stimulus sentence, is given in Appendix E. There, and hereafter, these tapes are referred to as "test tapes" for convenience.

Two practice sentences were recorded and prepared in like manner. These test tapes and practice tapes were played back to the subjects in the experimental session through the tape recorder variable delay system described below.

Experimental Apparatus

To implement the procedure for adjustment of pause duration, a modified two-track playback head system was constructed so that pause duration adjustment could be accomplished through delaying, to a greater or lesser degree, the playback of the Track B signal (that part of the sentence beginning with the word "contact") relative to the signal on Track A (that part of the sentence ending with the word "lost"). To provide for such a variable delay, a tape guide attached to a worm gear assembly, as shown in Figure 2, was mounted between the fixed Track A and Track B playback heads. Since the tape guide was attached to the worm gear, its position, relative to the horizontal

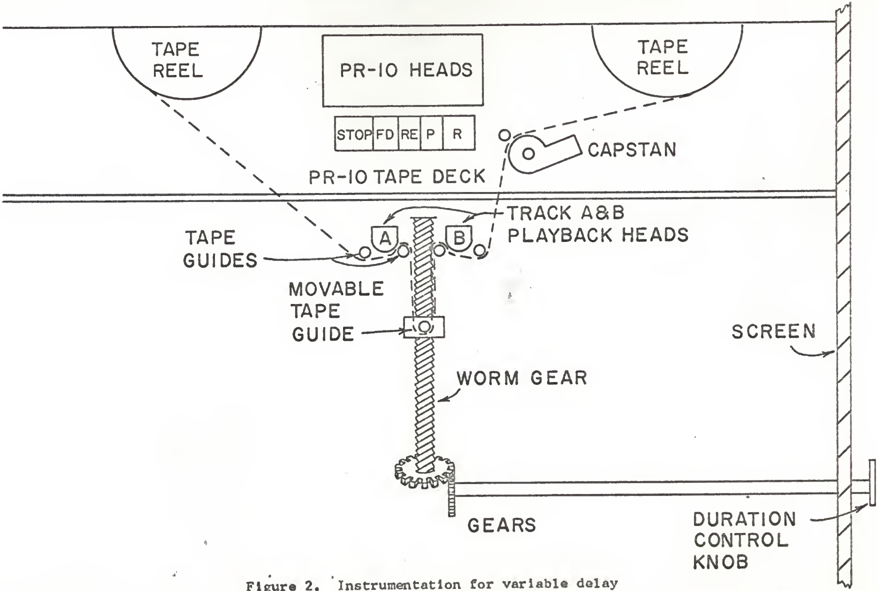

Figure 2. Instrumentation for variable delay system to adjust pause durations.

plane of the playback heads, could be varied up and down the worm gear, thereby increasing or decreasing the effective distance between the playback heads. Pause duration between "lost" and "contact" could be increased, then, by moving the tape guide away from the playback heads, since this movable guide altered the time of onset of the Track B signal relative to that on Track A by changing the distance the tape must travel to reach the "B" playback head.

This playback head assembly was coupled to an Ampex PR-10 tape deck and stereo amplifier assembly driving a pair of Sharpe HA-10 earphones with a monaural Y-cord input so that the Track A and B signal was received in both ears (see Figure 3).

To block this experimental apparatus and other instruments from the subjects' view, a large screen was placed between the subject and the equipment. The durational control knob protruded through this screen (see Figure 2) so that pause duration adjustments could be made by the subjects from that side of the screen opposite these instruments. To further eliminate visual cues from affecting a subject's adjustment of pause duration, the duration control knob was unmarked save to indicate direction to increase or decrease.

The outputs from Channel A and Channel B of the playback recorder were coupled respectively to Channel A and B inputs of a Tektronix 564 Storage Oscilloscope as depicted in Figure 3. This oscilloscope is a special purpose instrument designed to store cathode-ray tube displays for viewing or photographing. Each stimulus sentence, as adjusted by the subject, was viewed on the oscilloscope screen. When that portion of the signal containing the end of Channel A signal and the beginning of Channel B signal was

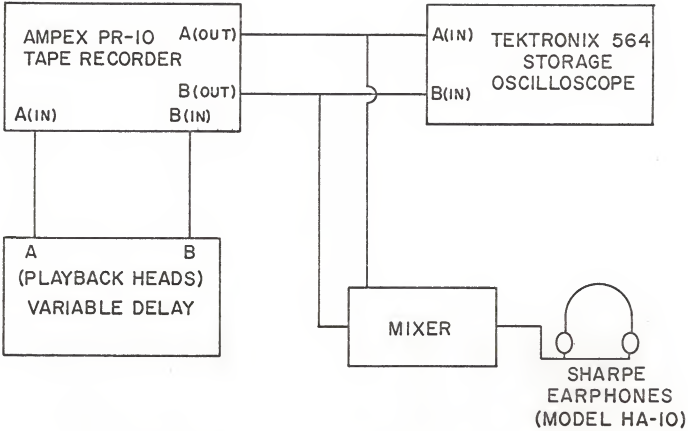

Figure 3. Block diagram of instrumentation for experimental pause adjustment and measurement.

located visually, it was stored on the oscilloscope display. A permanent record of this display was then obtained by photographing it with a Polaroid trace-recording camera attached to the face of the oscilloscope screen. These photographs were then put aside for later analysis.

Perceptual Tasks

A psychophysical procedure was employed wherein the subjects adjusted the duration of the variable silent interval (between the words "lost" and "contact") within each stimulus sentence for each of three perceptual tasks: (1) pause detection threshold, (2) fluent pause and (3) hesitation pause.

Each subject was instructed initially as to the general nature of the tasks required and presented with the following practice sentences for familiarization: (1) "They passed quickly through the state." and (2) "The day passed quickly." Practice was allowed on these sentences until both the subject and experimenter were satisfied that the instructions were understood. The practice sessions averaged approximately ten minutes per task.

In the detection task, the subjects were asked to adjust the duration of the silent interval between the words "lost" and "contact" to that point at which they just detected a pause. This adjustment began with the pause duration between "lost" and "contact" set at 500 msec. (arbitrary starting point). The subject was instructed to decrease the duration of this pause until he perceived a definite overlap of the words "lost" and "contact" and then increase the duration again until he just detected a pause. The subject was encouraged to make as many such "threshold crossings" as necessary in order to locate that point at which the duration of the silent interval was just sufficient to be perceived as a pause; that is, if the duration were any shorter, the interval would not be perceived as a pause. When the subject indicated he had located this point, the adjusted sentence was displayed on the oscilloscope screen and photographed for subsequent analysis.

For the task which required adjustment of the optimal duration of fluent pauses, as well as for the adjustment of hesitation pauses, the subject was merely informed that there are two types of pauses: (1) pauses which do interrupt the smooth flow of speech (hesitation pauses) and (2) pauses which do not interrupt the smooth flow of speech (fluent pauses). No more detailed description of these pauses was given (such as: fluent pauses are said to be used as oral punctuation, or, hesitation pauses are silent intervals of unusual length) since such specific information might have inadvertently influenced the subjects' performances.

For the fluent pause task, the subjects were presented a sentence in which the duration of the silent interval to be adjusted was arbitrarily set at 500 msec. They then were instructed to adjust the duration of this interval, either increasing or decreasing it, to the point where it was perceived as being an optimal duration, that is, the duration which was best suited for conveying the meaning of the sentence in a fluent, intelligible manner. Subjects were allowed as many trials as necessary in making this adjustment; when they indicated that the duration of the pause was optimal for the fluent presentation of the sentence, it was displayed on the oscilloscope screen and photographed.

The hesitation pause task proceeded exactly as the fluent pause task except that subjects were requested to adjust the pause duration to that point where they just perceived it as being a hesitation.

There were three replications of each task for each subject. Order of sentence presentation across the three replications within a task was completely randomized. Each task, including replications, was completed before going on to the next. The order in which subjects performed the three experimental tasks was likewise randomized across subjects.

Pause Duration Measurement

The adjusted pause durations were measured from the oscilloscope photographs in the following manner. Two judges were selected to measure independently the distance in centimeters between the end of the Channel A signal and the beginning of the Channel B signal for each of the photographs. Each judge was first familiarized with the prominent landmarks for each of the five stimulus sentences. This included particularly the shape and approximate location of the noise burst of the plosive /t/ which signaled the end of the Channel A signal for sentences two, three, four and five and the fact that there was no such visible noise burst in sentence one. The judges were asked to locate that point at which the Channel A signal reached the baseline and remained there. Likewise, they were instructed to find the point at which the integrated Channel B signal dropped from the baseline. Since a fading of the signal

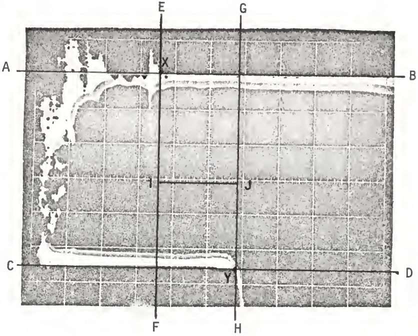

Figure 4. Reference points used in the measurement of pause durations from the oscillographic photographs: line AB, the top of the Channel A baseline; line EF, a perpendicular erected through AB at point X (point where the speech signal drops to the baseline AB); line CD, the bottom of the Channel B baseline; line GH, a perpendicular erected at point Y (point at which the Channel B integrated signal drops below the baseline); line IJ, the distance in centimeters from the end of Channel A signal to the beginning of Channel B signal.

often occurred in the center of the oscilloscope screen, the thickness of the baseline varied in this area of the picture which made identification of the end of Channel A signal and the beginning of Channel B signal difficult.

To afford a more stable criterion of what constituted the Channel A baseline then, the judges were instructed to lay a straight-edge ruler along the top of the baseline in the extreme right-hand corner[4] of the photo display so that it was parallel to the horizontal scale division. This procedure is depicted as line AB in Figure 4. Judges were then instructed to find that point, going from right to left along line AB, where the speech signal first departed from this baseline reference (indicated by point X in Figure 4) and draw a perpendicular (line EF) at this point. Likewise, they were asked to lay the straight-edge ruler along the bottom of the Channel B baseline in the extreme left-hand corner[4] of the photo display, parallel to the horizontal scale division (line CD), locate that point (Y) at which the signal dropped below this baseline reference and draw a perpendicular at this point also (line GH). The distance between lines EF and GH is depicted in Figure 4 as line IJ. This distance represents the time between the end of the Channel A signal and the beginning of the Channel B signal and is converted to pause time in msec. by multiplying this distance by the time scale value used for that particular oscillographic display.

---

[4] The top of the Channel A baseline as displayed in the upper right-hand corner of the display and the bottom of the Channel B baseline as it was displayed in the lower left-hand corner were chosen as baseline references since the least fading and distortion of the baseline signal occurred at these points.

Inter-judge agreement on these measures was computed in terms of the percentage of instances in which the measures of the two judges agreed within plus or minus 1 mm. Such agreement was reached in 465 of the 540 measures for 86 per cent agreement.

Most of the photographs on which the judges' measurements did not agree were of rather poor quality in that the oscilloscopic display had faded somewhat by the time it had been photographed. Since the two judges indicated that it was often difficult to identify the Channel A signal landmarks, particularly the plosive noise burst, because of this fading, they were given good quality oscillographic photographs of each of the five stimulus sentences to use as a visual reference to aid in the identification of the Channel A landmarks. They were then asked to remeasure, with the aid of these reference photographs, the 75 pause durations on which they had disagreed. The same measurement procedure as previously was utilized.

Upon remeasuring, agreement within plus or minus 1 mm. was reached in 38 of the 75 measurements. The remaining 37 measures were given to a third judge. His measurements agreed with one or the other of the original two judges in all 37 instances.

Statistical Analysis

The mean duration of the two agreeing judges' measurements for each of the 540 oscillographic displays was computed and used as the criterion measure of pause duration in the statistical analysis. For ease of computation in this analysis, a constant of 300 was added to each measure to eliminate any negative values.

A treatments-by-subjects analysis of variance was performed on the pause duration data of Appendix F to determine the effects of syntactic complexity, task and replications on pause durations (adopted from Lindquist, 1956). Error terms for this analysis were determined by deriving the expected mean squares and appropriate F-ratios as described by Winer (1962, p. 195).

In addition, two analyses of variance with repeated measures (Winer, 1962) were performed on the pause duration data to test for subject effects. The first of these analyses utilized the three replications as a within cell error term and the effect of subjects on pause duration adjustment was tested. In the second analysis, subjects were divided into two groups; those who were highly variable on the pure tone duration adjustment used in the screening task and those who had relatively less variation on the screening task. This analysis was performed to determine if there were any differences between these high and low variable groups.

CHAPTER III

RESULTS AND DISCUSSION

The purpose of this study was to provide information concerning the following questions: (1) Does duration of the silent interval differ for pause detection thresholds, fluent pauses and hesitation pauses within identical environments? (2) Does duration of the silent interval for pause detection threshold, fluent pause and hesitation pause vary as a function of syntactic complexity? (3) How variable are subjects' perceptions of the duration of silent intervals? The results are presented and discussed relative to each of these questions.

Comparison of Durations for Pause Detection Threshold, Fluent Pause and Hesitation Pause

The mean durations for pause detection threshold, fluent pause and hesitation pause are shown in Figure 5. From this figure, it can be seen that the mean hesitation pause duration (505 msec.) differs considerably from the mean fluent pause duration (186 msec.) and that both of these pause types differ in duration from the mean pause detection threshold (6 msec.). To determine whether these differences were statistically significant (as well as the differences between levels of syntactic complexity and replications), a treatment by treatment by treatment by subjects analysis of variance (adapted from the model provided by Lindquist, 1956) was performed.

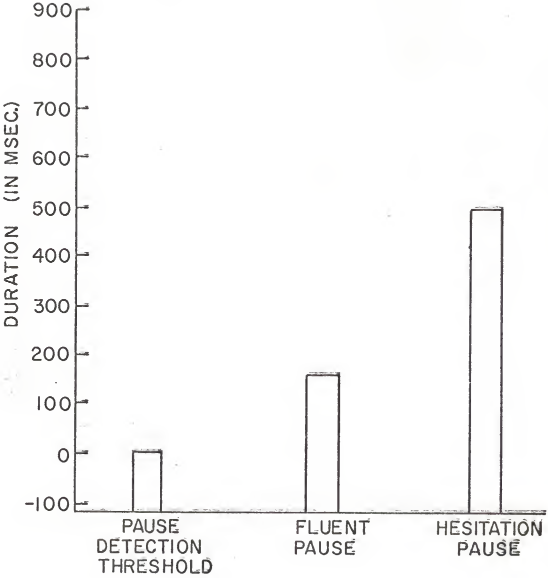

Figure 5. Comparison of mean durations of pause detection threshold, fluent pause and hesitation pause.

The results of this analysis, presented in Table 4, show that the effects of task, syntactic complexity and the task by syntactic complexity interaction were significant at the .05 level of confidence. The significant interaction was further analyzed by computing the simple main effects at each level of the other factor. The outcome, summarized in Tables 5 and 6, indicates no significant differences among the durations of pause detection threshold, fluent pause and hesitation pause at the two lowest levels of syntactic complexity; there are significant differences, however, at the other three levels. Since the syntactic boundaries at the two lowest levels of complexity represent within-phrase relations and those at the higher levels of complexity represent between-phrase relations, the effects of task on pause duration will be discussed, in turn, with respect to this within-phrase, between-phrase distinction.

Within-Phrase Pause Durations

The lack of significance among pause detection threshold, fluent pause and hesitation pause durations at the within-phrase pause boundaries in sentences one and two suggests that within-phrase pauses are functionally distinct from between-phrase pauses. Goldman-Eisler (1968), as well as Boomer and Dittman (1962), makes a similar distinction between juncture (grammatical) pauses and hesitation pauses in which the former are limited to between-phrase and the latter to within-phrase occurrences. It is clearly implied in their works that duration of the silent interval is the only significant cue for within-phrase hesitation pauses; the present study, however, indicates that duration, in itself, is not a sufficient cue for such pauses. An alternative is that allophonic changes, such as prepausal phoneme

Table 4. Summary of treatments by subjects analysis of variance to determine the effects of task, syntactic complexity and replications on pause durations.

Source of Variation	SS	df	MS	F
A (task)	22,932,549	2	11,466,275	43.63*
B (syntactic Complexity)	3,552,238	4	888,059	6.53*
C (replications)	99,528	2	49,764	4.65
S (subjects)	3,436,476	11	312,407	
AB	6,837,168	8	854,646	8.97*
AC	12,182	8	1,523	0.02
BC	69,403	4	17,351	1.34
AS	5,986,597	44	136,059	
BS	5,782,012	22	262,819	
CS	235,565	22	10,707	
ABC	61,430	16	3,839	0.05
ABS	8,385,745	88	95,293	
ACS	588,027	88	6,682	
BCS	569,359	44	12,940	
ABCS	1,483,014	176	8,426	
Total	60,031,294	539		

*Significant at .05 level of confidence (conservative test:[a] $F_{.05}$, 1, 11 = 4.84).

[a] It was not assumed in this study that the homogeneity assumptions underlying this statistical model (see Winer, 1962, pp. 290-292) were met. Therefore, a conservative test (Winer, 1962, p. 322) was used to test all main and interaction effects for this and all subsequent analyses.

Table 5. Simple main effects for factor A (tasks) at each level of factor B (syntactic complexity).

Source of Variation	SS	df	MS	F
A @ B_1	648,627	2	324,314	1.23
A @ B_2	1,019,351	2	509,675	1.94
A @ B_3	3,960,353	2	1,980,177	7.53*
A @ B_4	10,029,195	2	5,014,598	19.08*
A @ B_5	14,112,191	2	7,056,095	26.85*
Error (A @ Bj)	5,782,012	22	262,819	

*Significant at the .05 level of confidence (conservative test: $F_{.05}$, 1, 11 = 4.84).

Table 6. Simple main effects for factor B (syntactic complexity) at each level of factor A (tasks).

Source of Variation	SS	df	MS	F
B @ A_1	313,866	4	78,467	0.58
B @ A_2	418,271	4	104,568	0.77
B @ A_3	9,657,269	4	2,414,317	17.74*
Error (B @ Ai)	5,986,597	44	136,054	

*Significant at .05 level of confidence (conservative test: $F_{.05}$, 1, 11 = 4.84).

lengthening, combine with silent interval duration in providing for the perception of within-phrase hesitation pauses.

The discrepancy that has been cited might relate as well to definitional considerations. In the present study, hesitation pauses were defined with reference to fluent pause durations; it is not clear that this was the case in the earlier investigations. Although Maclay and Osgood (1959), Goldman-Eisler (1968) and Lounsbury (1954) define hesitation pauses as silent intervals of "unusual" duration, they fail to state an explicit reference for the concept "unusual." Their conception of "unusual" duration may not refer to fluent pause duration, but rather to some internalized standard of the listener. The large degree of variability between subjects' mean hesitation pause durations (see Appendix G for mean and standard deviation values for pause durations) suggests that this is plausible.

A further consideration concerns the problem of sampling in the present investigation, both in terms of number of sentences and subjects. It is apparent from looking at the mean duration values for sentences one and two (see Figure 6) that the mean hesitation pause duration of 265 msec. is considerably longer than the mean fluent pause duration (143 msec.) and pause detection threshold (52 msec.). It should be re-emphasized, however, that despite the magnitude of these differences, they were not statistically significant. More subjects and more sentences sampling the within-phrase pausal duration may lead to such statistically significant differences. If this did occur, the within-phrase, between-phrase distinction would be obviated. However, at present, this remains a topic for further study.

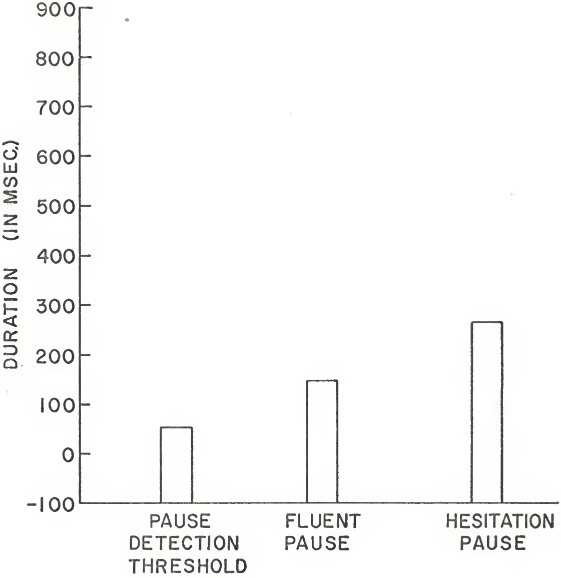

Figure 6. Comparison of sentence one and two mean pause durations for pause detection threshold, fluent pause and hesitation pause.

Between-Phrase Pause Durations

To determine which of the tasks accounted for the overall significance of this effect in the between-phrase pause durations (sentences three, four and five), the Newman-Keuls test of treatment differences following an overall significant F-ratio[5] was applied to this portion of the data. The results of this test presented in Tables 7, 8 and 9 indicate that, in sentence three, hesitation pause duration is significantly different from both fluent pause duration and pause detection threshold, but the latter two do not differ significantly from each other. This failure of the difference between fluent pause duration and pause detection threshold to reach statistical significance in sentence three should be viewed cautiously since it does approach the critical value for the .05 level of confidence. However, pause durations for pause detection threshold, fluent pause and hesitation pause do differ significantly from each other in sentences four and five.

Essentially, then, the data for between-phrase pauses can be viewed as support for the notion that pause detection threshold, fluent pause and hesitation pause are perceptually differentiated on the basis of duration of the silent interval. However, it should be noted that in many previous investigations, between-phrase pauses have been systematically excluded. This has been particularly true in the study of hesitation pauses (Goldman-Eisler, 1968; Boomer and Dittman, 1962), although the rationale for the exclusion of such

---

[5] This statistic tests the difference between any number of means arranged in increasing order of magnitude while maintaining the level of significance equal to alpha.

Table 7. Comparison of between-phrase pause durations at sentence three for pause detection threshold, fluent pause and hesitation pause. The marginal values are the ordered pause duration means. The values in the matrix represent the differences between each marginal pair.

	Pause Detection Threshold -26	Fluent Pause 167	Hesitation Pause 441
Pause Detection Threshold -26	-	193	467*
Fluent Pause 167		-	274*
Hesitation Pause 441			-

*Significant at .05 level of confidence. Critical values: $W_2 = 206$; $W_3 = 253$.[b]

[b] $W_r = q_{.05}, r, v(MS_{as}/n)^{\frac{1}{2}}$ where q is the Studentized Range Statistic, .05 represents the confidence level, r is the number of steps between means and v refers to the degrees of freedom for the error term, MS_{as}.

Table 8. Comparison of the between-phrase pause durations at sentence four for pause detection threshold, fluent pause and hesitation pause. The marginal values are the ordered means. The values in the matrix represent the differences between each marginal pair.

		Pause Detection Threshold -5	Fluent Pause 270	Hesitation Pause 734
Pause Detection Threshold	-5	-	275*	739*
Fluent Pause	270		-	464*
Hesitation Pause	734			-

*Significant at .05 level of confidence. Critical values: $W_2 = 206$; $W_3 = 253$.

Table 9. Comparison of the between-phrase pause durations at sentence five for pause detection threshold, fluent pause and hesitation pause. The marginal values are the ordered pause duration means. The values in the matrix represent the differences between each marginal pair.

		Pause Detection Threshold -44	Fluent Pause 206	Hesitation Pause 817
Pause Detection Threshold	-44	-	250*	861*
Fluent Pause	206		-	611*
Hesitation Pause	817			-

*Significant at .05 level of confidence. Critical values: $W_2 = 206$; $W_3 = 253$.

between-phrase pauses was never made clear. While the present data do suggest that within-phrase pauses are functionally distinct from between-phrase pauses, they would also argue against any such exclusion. For instance, if the present study had been confined to either within- or between-phrase pauses, the answer to the question concerning the role of silent interval duration in the perception of pauses would have been answered quite differently. It is apparent, therefore, that an understanding of the perceptual nature of speech pauses, per se, must take into account both within-phrase and between-phrase environments if it is to be considered at all valid and meaningful.

The question initially posed as to the extent to which pause duration is used to differentiate pause detection thresholds, fluent pauses and hesitation pauses may, therefore, be answered briefly as follows. Pause duration was found to be an insufficient perceptual cue, in and of itself, in differentiating between these pause types when such pauses occurred within phrases. However, with regard to between-phrase pauses, duration of the silent interval was a significant perceptual cue. In light of the differences found in the perception of within- and between-phrase pauses, further explication of the perceptual nature of speech pauses will have to take this within-between phrase distinction into account. Therefore, these results suggest that for within-phrase pauses, at least, other perceptual cues, as well as duration of the silent intervals, should be considered in any study of the differences between fluent and hesitation pauses.

Effects of Syntactic Complexity on Pause Durations

The second major question this study sought to explicate concerned effects of syntactic complexity on duration of pause detection threshold, fluent pause and hesitation pause. The overall mean pause duration for each level of syntactic complexity is shown in Figure 7. As mentioned previously, the treatments by subjects analysis of variance (see Table 4) indicates that the main effect of syntactic complexity is statistically significant. However, since there is also a significant interaction between tasks and levels of syntactic complexity, a simple main effects test was performed for syntactic complexity at each level of tasks.

The results of this analysis, presented in Table 6, indicate that level of syntactic complexity is a significant factor only for the hesitation pause task. Figure 8 clearly illustrates that as the level of syntactic complexity increases from levels two to five, the duration of the hesitation pause similarly increases. The increase in hesitation pause duration from syntactic complexity level one to level two, however, is slight.

To determine which of these levels of complexity was contributing to the overall simple main effect, a Newman-Keuls test of treatment differences was applied to the data. The results of this analysis, presented in Table 10, indicate that mean hesitation pause durations at levels four and five differ significantly from the mean hesitation pause durations at levels one, two and three. Hesitation pause durations at levels four and five, however, do not differ significantly, nor are there any significant differences in the durations among levels one, two and three.

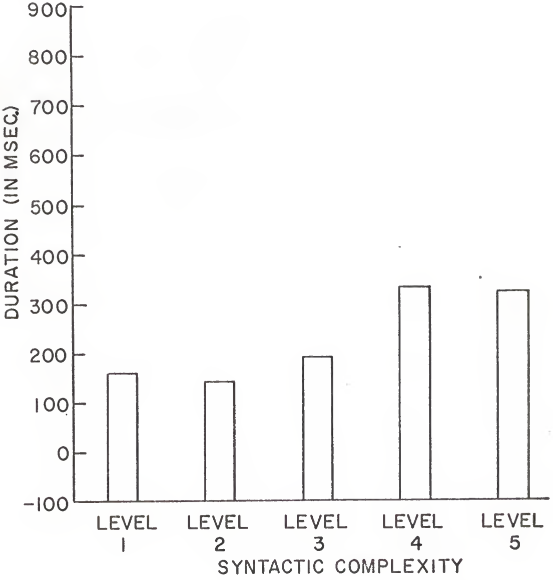

Figure 7. Mean pause duration for each of five levels of increasing syntactic complexity.

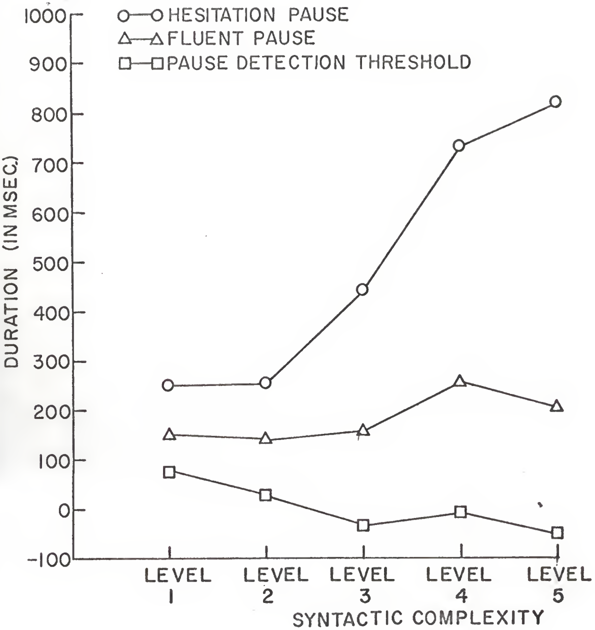

Figure 8. Mean pause duration of the five levels of increasing syntactic complexity for each task.

Table 10. Comparison of hesitation pause duration means for each level of syntactic complexity. The marginal values contain the ordered mean durations for each level of complexity. The values in the matrix represent the differences between each marginal pair.

	Level 1 262	Level 2 270	Level 3 441	Level 4 734	Level 5 817
Level 1 262	-	8	179	472*	555*
Level 2 270		-	171	464*	547*
Level 3 441			-	293*	376*
Level 4 734				-	83
Level 5 817					-

*Significant at the .05 level of confidence. Critical values: $W_2 = 191$; $W_3 = 262$; $W_5 = 281$.

These results can be interpreted as suggesting that syntactic complexity as measured by the two indices, structural complexity index and subjects' rankings of syntactic complexity, does not have an effect on pause durations. While the syntactic complexity of the pause boundaries in sentences four and five differed considerably on both complexity indices (see Tables 2 and 3), hesitation pause duration did not differ significantly at these boundaries. Sentences one, two and three are similarly different from each other in terms of the complexity indices and likewise show no significant differences in hesitation pause duration among themselves. In this respect, the pause boundaries in sentences four and five are alike in that both are subordinate clause-main clause boundaries, while those in sentences one, two and three are other types of syntactic boundaries (i.e., verb-object boundaries, adjective-noun boundaries, etc.). For convenience, the former will henceforth be referred to as subordinate clause boundaries and the latter simply as other boundaries.

More specific implications of these results for pause detection threshold, fluent pause and hesitation pause, in turn, are presented in the following discussion.

Pause Detection as a Function of Syntactic Complexity

As mentioned previously, syntactic complexity had no significant effect on pause detection threshold. From inspection of Figure 8, however, it is evident that there is a tendency for the pause detection threshold to decrease as syntactic complexity increases. Moreover, this decrease resulted in negative pause durations (that is, the point at which the subject indicated he just detected a pause was in reality an overlap of the words "lost" and "contact"). Inspection of the raw data (Appendix F), however, indicated that of the 36 data points for syntactic complexity level one, there was only one negative pause detection threshold. This was considerably fewer than occurred at the other four levels.

Comparison of the acoustic signals of the five stimulus sentences did reveal a possible explanation for these negative values. In sentences two through five, the word "lost" was terminated by a plosive burst which was perceptible both in the auditory and visual (oscilloscope) signal. Sentence one contained no such plosive burst. Thus, it was hypothesized that the subjects were disregarding, or were perceptually unaware of the overlapping of this plosive burst with the word "contact." This plosive burst overlap with the word "contact" may well have not been perceived since the plosive burst is often redundant information for the identification of plosive stops (Cooper, Delattre, Liberman, Borst and Gerstman, 1952). Therefore, the silent interval corresponding to the period of closure for the stop consonant may well have been taken to be part of the pause interval.

If it were assumed that the subjects did indeed disregard the plosive burst in the pause detection task, then a pause duration remeasurement[6] which also ignored this plosive burst should eliminate the negative pause detection thresholds. Such was clearly the case as Figure 9 illustrates. Not only did this remeasurement eliminate the negative pause detection thresholds, but it also made the pause detection curve resemble the curve for fluent pauses. This striking similarity between these two curves would suggest, then, that fluent pause duration at a particular boundary may well be a function of the pause detection threshold at that same boundary. This, however, needs verification by further research.

---

[6] The same three judges utilized in making the initial measurements also made these remeasurements. Instructions for this remeasurement differed from the original only in that judges were told to ignore the plosive burst as a part of the speech signal.

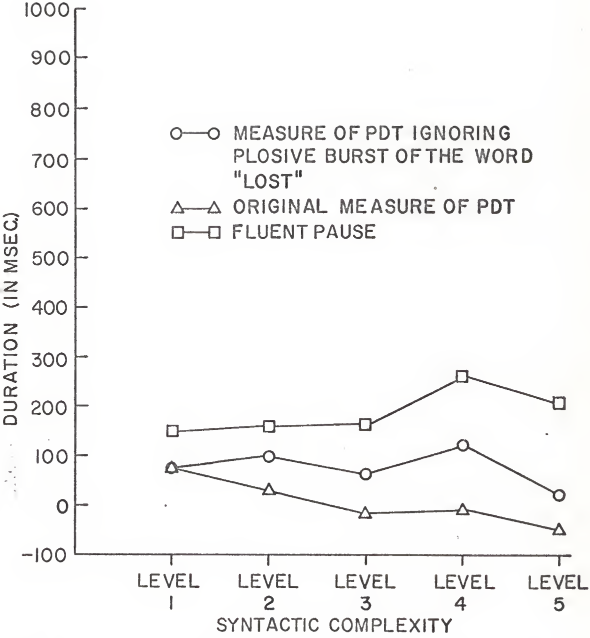

Figure 9. A comparison of fluent pause durations with pause detection thresholds obtained from measures in which the plosive burst terminating the word "lost" was included as part of the speech signal and pause detection thresholds obtained from measures in which this plosive burst was ignored.

The remeasuring of the pause detection thresholds also had the effect of increasing the overall mean pause detection threshold from 6 msec. to 77 msec. This value, however, is still far short of the minimum detected pause duration of 190 msec. reported by Agnello (1963). There are, however, several important procedural differences which might account for this discrepancy. First of all, the present study directed the subjects specifically to locate the duration at which the silent interval was just detected as a pause. Agnello assumed, *a priori*, that pauses shorter than 150 msec. were not perceived. Furthermore, the pause detection thresholds reported in this study are the result of direct measures, whereas Agnello's minimum pause duration is merely an estimate based on comparison of the total number of perceived pauses with the total number of pauses measured by machine (Agnello, 1963).

Agnello's data cannot, therefore, really be taken as an indication of the pause detection threshold in speech, nor can they really be compared to the results of the present study. The present data, likewise, would hardly suffice as anything more than an indication of the speech pause detection threshold. Considerably more data, obtained from more subjects and more phonetic environments, are needed before any substantial statements about such thresholds can be made.

Fluent Pause Duration as a Function of Syntactic Complexity

The finding that fluent pause duration did not differ significantly across the five levels of syntactic complexity has import in regard to the assumption that fluent pauses are used to signal the syntactic structure in speech (Carrell and Tiffany, 1960; Jones, 1962; and

Lieberman, 1967). On the basis of these results, one would have to conclude that this is not one of the primary functions of fluent pauses. This has particular relevance to Lieberman's (1967) notion that fluent pause duration is a primary cue marking syntactic junctures, or disjunctures as he refers to them, in speech. If his notion were, in general, correct, one would expect a difference in fluent pause durations at least for the between-phrase and within-phrase distinctions. Since such was not the case, the data from the present study are seen as contradictory evidence for the claim that fluent pause duration is a primary cue for disjuncture. On the other hand, the data can be interpreted as direct support for the interpretation advanced by Scholes (1968) that fluent pause duration is a relatively unimportant physical cue for signaling syntactic junctures (disjunctures).

This is not to say that fluent pause duration may not be used as a cue for disjuncture. The data merely suggest that, as a general rule, duration of the silent interval is not utilized for this purpose. It may also well be that, given specific instructions to emphasize or de-emphasize a particular part of a sentence, pause duration will be a primary parameter employed. Such a function of fluent pauses is suggested by Carrell and Tiffany (1960).

Then again, fluent pause duration may serve no other purpose in speech than to mark perceptual units in speech and to provide perceptual processing time as suggested by Aaronson and Markowitz (1967). These are only speculations, however, and must be researched further. The data from the present study offer little insight with regard to these possible functions of fluent pauses.

Hesitation Pause as a Function of Syntactic Complexity

A comparison among mean hesitation pause durations at the five levels of syntactic complexity, summarized in Table 10, shows clearly that the most complex syntactic relations (levels four and five) are associated with significantly longer hesitation pause durations than are the less complex relations (levels one, two and three). As mentioned previously, the variable pause boundaries in sentences four and five differ from those in sentences one through three in that the former are subordinate clause boundaries while the latter are other types of syntactic boundaries. Since subordinate clause boundaries are frequently punctuated by a comma in orthography, these results can be viewed as being particularly relevant to the notion that fluent pauses in speech are often used as oral punctuation (Carrell and Tiffany, 1960). Although the comparison of fluent pause durations did not reveal a significant differential use of pause durations between subordinate clause and other syntactic boundaries, the use of hesitation pause durations in this regard has some interesting implications.

The major implication of this finding is that, although the optimal fluent pause duration between subordinate clause and other boundaries does not differ significantly, the range of fluent pause durations (that range of durations extending from the pause detection threshold to the minimum hesitation pause duration) is greater at the subordinate clause boundaries (917 msec.) than at the other boundaries (496 msec.). Although this range may not be utilized extensively, its potential is there as can be evidenced by this difference.

An equally important finding is that the within-phrase, between-phrase distinction frequently employed in the study of hesitation pauses (Maclay and Osgood, 1959; Goldman-Eisler, 1968) may not be as significant as commonly thought. Goldman-Eisler (1968), in particular, frequently limits her study of hesitation pauses to within-phrase occurrences such as those corresponding to syntactic complexity levels one and two of this study. In this regard, it may be interesting to note that the mean within-phrase pause duration in the present study of 265 msec. compares favorably with the 250 msec. standard adopted by Goldman-Eisler (1961b) and to a lesser degree, the 200 msec. "hesitation pause threshold" reported by Boomer and Dittman (1962) at within-phrase boundaries. However, the data in the present study also indicate that the hesitation pause durations at the syntactic complexity level three are not significantly different from those at levels one and two, yet the former is a between-phrase boundary and the latter are within-phrase pause boundaries. This suggests that any generalization about hesitation pauses based largely on within-phrase occurrences is of questionable validity. Furthermore, the differences between hesitation pause durations at subordinate clause boundaries and other boundaries argue rather forcibly against the notion that a hesitation pause threshold can be established irrespective of environment and/or perceptual judgments as has been commonly assumed (Goldman-Eisler, 1961a, 1961b, 1968; Boomer and Dittman, 1962).

The distinctions of within-phrase and between-phrase pauses and the subordinate clause and non-subordinate clause pauses may, in this regard, be further useful to demonstrate the behavioral relevance of these units. That is, the differential distribution of hesitation pauses at the subordinate clause and other types of syntactic boundaries can be interpreted as evidence supporting the functional distinction between such boundaries. Likewise, the finding that hesitation and fluent pause durations differed at between-phrase boundaries and not at within-phrase boundaries can also be interpreted to some degree as providing evidence for the psychological reality of the phrase.

Subject Variability

The third question raised in this study concerned subject variability in the perception of pause durations. Since it is possible to speak of two sources of subject variation, within and between, measures reflecting both types of subject variation were obtained and subjected to statistical analysis.

To provide an estimate of within subject variability, each subject was asked to replicate each pause adjustment three times. The overall mean pause duration for each of these replications is presented in Figure 10, and the replication mean duration for each task at each level of syntactic complexity is shown in Figure 11. It is evident that there is a slight overall increase in mean pause duration from one replication to the next, particularly in the hesitation task. The treatments by subjects analysis of variance, summarized in Table 4, indicated that these differences were not statistically significant. This suggests that, within themselves, the subjects were performing consistently from one replication to the next.

Since the replications effect was shown to be non-significant in the overall analysis of variance, the replication scores were pooled to provide an estimate of within cell variance to be used in

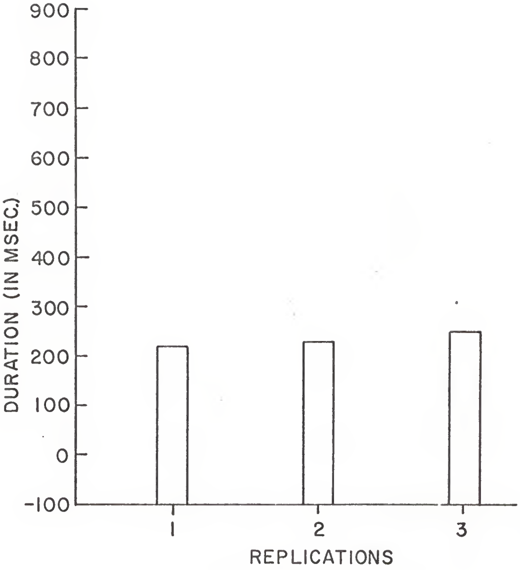

Figure 10. Overall mean pause duration for each of three replications.

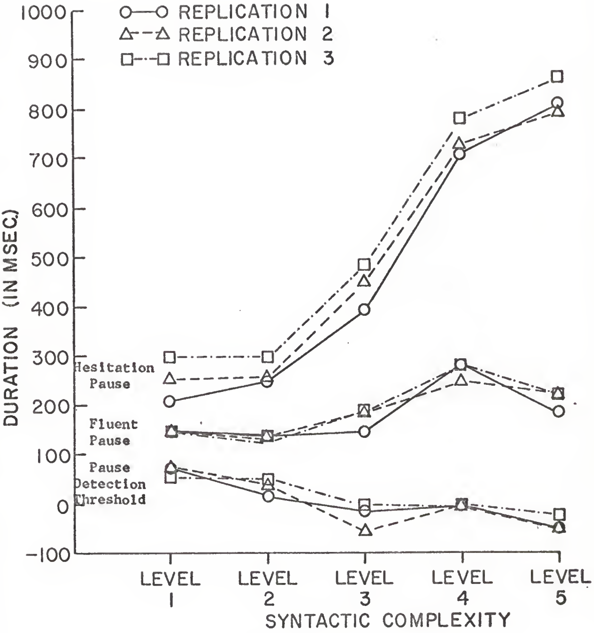

Figure 11. Mean pause duration for replications and tasks at each level of syntactic complexity.

a test of between subject differences. The results of this analysis (Table 11) show that subjects do differ significantly among themselves. Subject interactions with other factors could not be tested since the error term ordinarily used to test these interactions would have represented, in this instance, a pooling of heterogenous sources of variance as indicated by the F-max test performed on each of these error terms. Other main effects and interactions had been tested previously (Table 4) and were therefore not included in this analysis.

These subject differences in pause perception parallel the situation found in the production aspect. Goldman-Eisler (1961a, 1961b, 1968) and Scholes (1968) report considerable variation among individuals in the use of pauses in speech. The subject differences reported here in this perceptual study, as well as the reported individual variation in the production studies, thus point out one important fact. Any study of pauses, be it perceptual or production, should take care to use a relatively large group of randomly selected subjects since the outcome could well reflect the bias of a small group of subjects and not be representative of the population as a whole. Although the present study did restrict itself to a select group of subjects (male undergraduate college students), this was done to control for possible unwanted sources of variation due to sex, age, socioeconomic and/or educational differences. While there is no a priori reason to assume that these factors would affect the perception of pauses in speech, they do need to be considered for further study.

The fact that considerable individual differences exist both in the production and perceptual aspects of pausal behavior leads to

Table 11. Summary of the analysis of variance used to test for between subject effects.

Source of Variation	SS	df	MS	F
Between	3,771,569	35		
Subjects	3,436,475	11	312,406	22.38*
Within cell error	335,094	24	13,962	
Within	56,259,725	504	Not tested	
Total	60,031,294	539		

*Significant at the .05 level of confidence (conservative test: $F_{.05, 1, 2} = 18.5$).

the obvious conclusion that these individual differences may be related. More specifically, the question could be asked whether or not a person who habitually produces longer pauses at a particular pause boundary would also use a longer adjusted pause duration at this boundary in a perceptual study such as this one. It was not the intent in the present study to compare these comprehension and production aspects; it is, however, a prime topic for further study.

Data from the present study, however, were available which could be examined with respect to further explication of the question of subject variability in the perception of pauses. Since it was noted in the screening session that some subjects had been quite variable in making the tone duration match across the three trials, it was decided to divide subjects into two groups on the basis of high or low variability on the screening task. The data were then re-analyzed using a three-factor analysis of variance with repeated measures on two factors (Winer, 1962) to see if there were a differential performance between the two groups. Since replications had already been shown not to be a significant factor (see Table 4), the mean of the three replications was taken as the criterion measure in this analysis for ease of computation.

The results of this analysis, presented in Table 12, indicate no difference on the experimental tasks between the high and low variable screening task groups. The F-ratios for task, syntax and task by syntax interaction, however, were all significant at the .05 level of confidence.[7]

---

[7] These results substantiate those of the original analysis summarized in Table 4. The slightly smaller F-ratios in this analysis are primarily due to the error introduced by using the mean of the three replications as the criterion measure.

Table 12. Summary of analysis of variance used to test for differences between high and low variable groups of subjects.

Source of Variation	SS	df	MS	F
Between	10,308,864	11		
G (groups)	254,720	1	254,720	0.25
SWG	10,054,144	10	1,005,414	
Within	160,426,752	168		
A (tasks)	68,797,440	2	34,398,720	39.86*
AG	88,064	2	44,032	0.05
A X SWG	17,257,728	20	862,886	
B (syntactic complexity)	10,656,768	4	2,664,192	6.32*
BG	1,088,256	4	272,064	0.65
B X SWG	16,870,912	40	4,217,730	
AB	20,511,488	8	2,563,936	8.50*
ABG	1,020,416	8	127,552	0.42
AB X SWG	24,135,680	80	301,696	
Total	170,735,616	179		

*Significant at .05 level of confidence (conservative test: $F_{.05, 1, 10} = 4.96$).

The fact that there were no significant differences between the two groups can more than likely be attributed to the fact that the screening task and experimental tasks are quite dissimilar in several respects. The major point to be made in this regard is that the screening task required that each subject adjust the test tone duration to match that of an explicit reference tone. In the experimental tasks, if a reference were used at all, it was an internal reference and not explicit. The most obvious difference in the two tasks, however, is that one required the perception of filled intervals (pure tones) while the other required the perception of unfilled intervals. Steiner (1964) presents data which indicate that filled and unfilled durations are not perceived alike. That fact alone could account for the apparent unrelated nature of the screening and experimental tasks.

On the basis of the overall results, then, it can be stated that subjects are consistent in their perceptual judgments of speech pauses from one time to the next. However, there is considerable variability among subjects. This variability in the perceptual processing of speech pauses parallels the situation reported by studies in the production aspect of pauses also. Finally, degree of variability in making temporal matches with pure tones seems unrelated to subjects' performances in the perceptual processing of speech pauses.

CHAPTER IV

SUMMARY AND CONCLUSIONS

Previous studies of pausal behavior in speech have yielded inconsistent results concerning the role of silent interval duration in the perception of speech pauses. The present study was designed to investigate the relationship between syntactic environment and silent interval duration as these two factors affect the perception of pauses in speech. It was hoped, thereby, to provide some measure of specificity to this perplexing problem.

The psychophysical method of adjustment was employed to obtain the perceptual judgments of pause duration. These judgments were obtained within identical phonetic and syntactic environments for each of five recorded sentences. The sequence "lost contact" was embedded in each of these sentences in such a manner that the syntactic relations between "lost" and "contact" varied from sentence to sentence. These syntactic relations were specified both through use of a structural complexity index (Miller and Chomsky, 1963) and subjects' rankings of the complexity of these relations.

Pause adjustment was accomplished in the following manner. The five stimulus sentences were recorded by a single speaker who had been selected from an original pool of four on the basis of his high ranking on the six speaker attributes of quality, fluency, naturalness, rate, precision and overall effectiveness. These sentences were then dubbed onto a two-track tape such that the portion of the sentence up to and including the word "lost" was recorded on Track A, while the remainder of the sentence (beginning with the word "contact") was recorded on Track B. Pause duration between the words "lost" and "contact" could then be manipulated by mechanically delaying, to a greater or lesser degree, the playback of the Track B signal with respect to the Track A signal.

Twelve young adult male college students were selected as subjects for the study. Each subject, working individually, was asked to adjust the duration of the pause between the words "lost" and "contact" so that (1) the duration of the silent interval was just perceived as a pause, (2) the pause duration was considered optimal for the "fluent" presentation of the sentence and (3) the pause was just perceived as a hesitation. Each pause adjustment for the five sentences within each task was repeated three times to obtain an estimate of within subject variability. Subjects were given as much time and as many trials as necessary to make each adjustment. The interval of the sentence containing the adjusted pause was displayed on an oscilloscope screen and photographed. Pause durations were measured from these photographs.

The results of this study are best summarized with respect to the three major questions this study sought to answer. The first of these concerned the extent to which duration of the silent interval served as a cue for the perceptual differentiation of pause detection threshold, fluent pause and hesitation pause. The results show that for sentences one and two, duration of the silent interval for these three pause types did not differ significantly. For sentences three, four and five, however, the durations did differ significantly from each other. Since the adjustable pauses in sentences one and two can be characterized as within-phrase pauses and those in sentences three, four and five as between-phrase pauses, these results were interpreted generally as follows. Duration of the silent interval is a significant perceptual cue for the differentiation of pause detection threshold, fluent pause and hesitation pause only when such pauses are perceived in a between-phrase environment. For within-phrase pauses, however, it is apparent that duration of the silent interval, in and of itself, is not an adequate cue for the perceptual differentiation of these pause types as commonly thought (Goldman-Eisler, 1968; Boomer and Dittman, 1962; and Lounsbury, 1954). Other perceptual cues, such as pre-pausal allophonic changes, must be postulated as contributing to this overall differentiation.

Furthermore, the fact that pause duration seems to be an adequate perceptual cue for between-phrase pauses and an insufficient cue for within-phrase pauses demonstrates that pausal behavior cannot be realistically studied only in the within-phrase environment as has been done in the past (Goldman-Eisler, 1968). Stated differently, these data suggest that syntactic (within phrase or between phrase) environment must be considered in any comparison of the durations of these three pause types if the generalizations arising from these comparisons are to be considered valid.

This, then, leads to the second major question posed in this study which concerned the effects of syntactic complexity on pause durations within each of the three tasks. The results indicate that silent interval duration did not differ significantly from one level of syntactic complexity to the next for either pause detection threshold or fluent pause. Hesitation pause duration, however, did differ significantly as a function of syntactic structure in that hesitation pauses perceived within a subordinate clause-main clause boundary were significantly longer in duration than were those occurring in other types of syntactic boundaries.

On the basis of these results, therefore, it can be concluded that syntactic structure has no effect on the pause detection threshold. The results indicating that fluent pause durations are likewise not significantly different at the five levels of syntactic complexity can be viewed as a contradiction of the Lieberman (1967) claim that fluent pause duration is a significant cue for disjuncture. However, while the data indicate that fluent pause durations are not ordinarily "used" to signal syntactic structure, this is not to say that the capability does not exist to punctuate a sentence or to emphasize a portion of it by using fluent pause duration.

The data which show longer hesitation pause duration at subordinate clause boundaries can be construed, in this regard, as showing that the subject has at his disposal a wide range of fluent pause durations (extending from the pause detection threshold to the minimum hesitation pause duration) should he desire or need to "punctuate" the sentence at such boundaries. In fact, one of the major considerations of the finding that hesitation pause durations were significantly longer at subordinate clause boundaries concerns the implication that the range of fluent pauses is likewise greater at these boundaries.

A further consideration of the greater hesitation pause durations at the subordinate clause boundaries suggests that studies such as those by Goldman-Eisler (1961a and 1961b) in which generalizations about hesitation pauses were based largely on within-phrase (within clause) occurrences, are of questionable validity.

With regard to the third question posed in this study, the data show that subjects are consistent in their perceptual judgments of speech pauses from one time to the next. Among individual subjects, however, there exists a considerable degree of variability. This variability in the perceptual processing of speech pauses parallels the situation reported in the production aspect of pauses also (Goldman-Eisler, 1968). Because of these significant individual differences in the production and perceptual processing of speech pauses, then, it seems important that any rigorous study of speech pauses take this factor into account, at least with regard to subject selection.

Finally, aside from these major findings, the data from the present study can also be interpreted as providing at least indirect evidence for the behavioral relevance of some basic linguistic units. The differential distribution of hesitation pause durations may be taken as evidence supporting the functional distinction between the subordinate clause boundaries and other types of syntactic boundaries while the difference between within- and between-phrase pauses may likewise be taken as some degree of evidence supporting the psychological reality of the phrase.

APPENDICES

APPENDIX A

INSTRUMENTATION AND PROCEDURE USED IN THE DURATION MATCHING TASK OF THE SUBJECT SCREENING TEST

Each individual participating in the duration matching portion of the screening session was required to adjust the duration of a test tone of 250 Hz so that it matched the duration of a standard reference (also 250 Hz) of 400 msec. duration. The test tone, a continuous 250 Hz tone from a Hewlett Packard 202CR pure tone oscillator, was transmitted to the subject's earphones via a Grason Stadler 829E electronic switch (see Figure 12). The individual heard this tone (set at a "comfortable" intensity level) as long as the electronic switch was in its "on" state and ceased to hear it when the switch was "off."

This "on" and "off" state of the switch was controlled by a set of clicks (negative pulses) recorded on a two-track tape and played back through a two-channel tape recorder delay system (see the experimental apparatus section for a detailed description of this system). The Track A click served as the "on" trigger and the Track B click as the "off" trigger for the electronic switch (see Figure 12). The duration of the test tone in the subject's earphones was therefore dependent upon the duration between the Track A and B clicks. By effectively increasing or decreasing the duration between the Track A

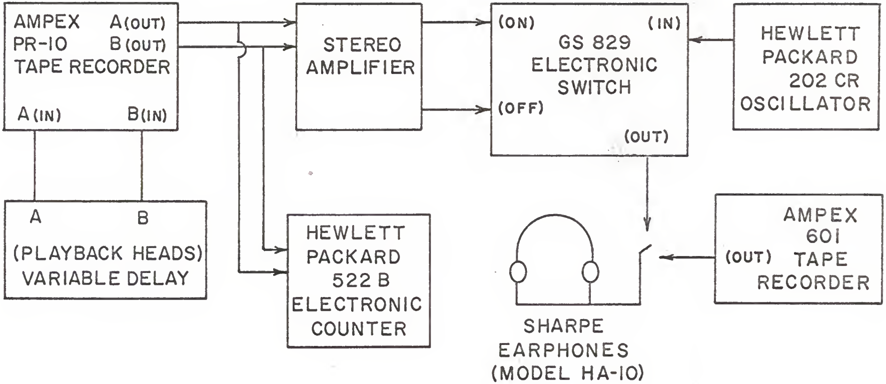

Figure 12. Block diagram of instrumentation used in the duration matching task of the subject screening test.

click and the Track B click, the duration of the test tone received by the subject was correspondingly increased or decreased.

The standard reference pure tone of 400 msec. duration was recorded repeatedly on a 1200 foot reel of tape with a five-second interstimulus interval. This tape was played back through an Ampex 601 tape recorder into the subject's earphones. The subject was able to listen to either the test tone or the standard tone (but not both at the same time) through the same set of earphones by selecting the desired signal with a switch to which the earphones were attached. Both the test and standard tones were presented binaurally.

The subject was thus asked to adjust the duration of the test tone to that point where he felt the duration was equal to that of the standard tone. The subject was given as much time as needed to make this match. When he indicated that the match had been made, the duration between the Track A click and Track B click was measured from the Hewlett Packard 522B electronic counter. This measure was then taken as the adjusted test tone duration. Three such duration matches were made by each subject and their respective durations obtained.

APPENDIX B

TREE DIAGRAM CONSTITUENT STRUCTURE ANALYSIS OF THE FIVE STIMULUS SENTENCES

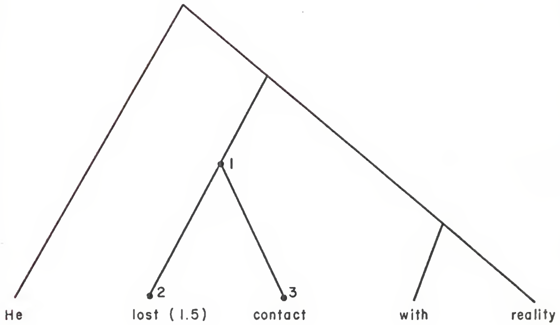

Figure 13. Tree diagram analysis of stimulus sentence one. The node-to-terminal-node ratio specifying the complexity of the relation between "lost" and "contact" is 3/2 (1.5).

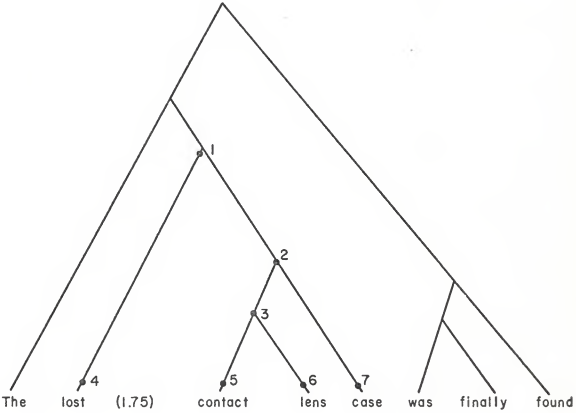

Figure 14. Tree diagram analysis of stimulus sentence two. The node-to-terminal-node ratio specifying the complexity of the relation between "lost" and "contact" is 7/4 (1.75).

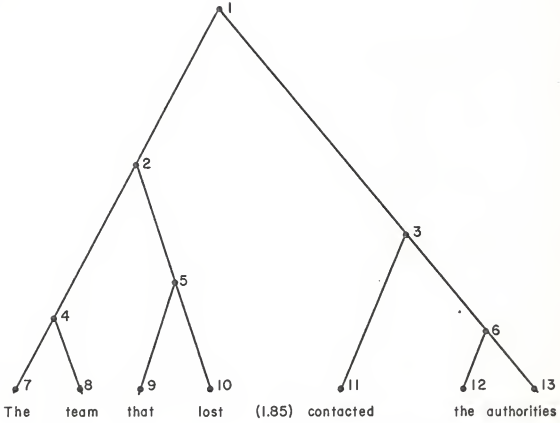

Figure 15. Tree diagram analysis of stimulus sentence three. The node-to-terminal-node ratio specifying the complexity of the relation between "lost" and "contact" is 13/7 (1.85).

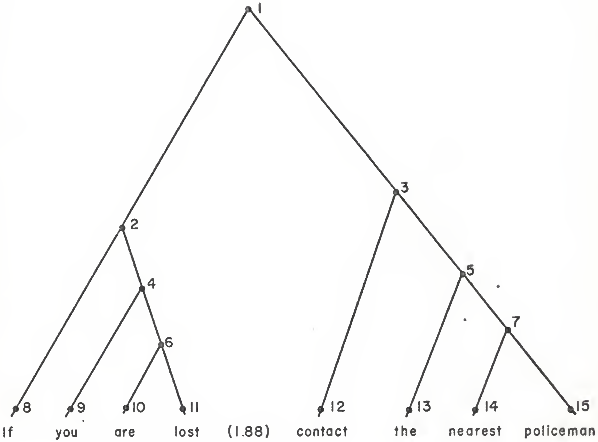

Figure 16. Tree diagram analysis of stimulus sentence four. The node-to-terminal-node ratio specifying the complexity of the relation between "lost" and "contact" is 15/8 (1.88).

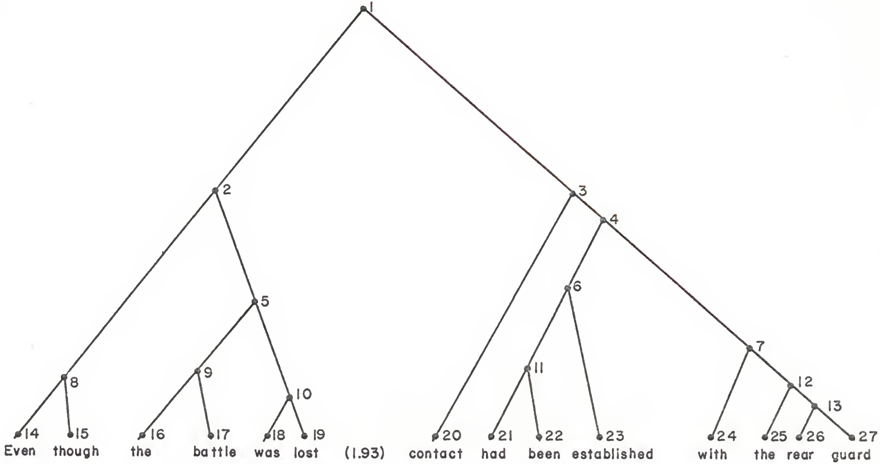

Figure 17. Tree diagram analysis of stimulus sentence five. The node-to-terminal-node ratio specifying the complexity of the relation between "lost" and "contact" is 27/14 (1.93).

APPENDIX C

INSTRUCTIONS TO SUBJECTS AND SENTENCES USED IN THE SYNTACTIC COMPLEXITY RANKING TASK

Instructions

As you know, the individual words of a sentence stand in some sort of syntactic relation to one another. For instance, in the sentence, "The brown leaves signal the approach of winter," the words "brown," "leaves," and "signal" enter into certain types of relations with one another. It should be noted, however, that these grammatical relations between the words or groups of words in a sentence can change rather markedly when they occur in another sentence. For instance, in the sentence, "Fred Brown leaves signal flares in the road," the sequence of words, "Brown," "leaves," and "signal," is exactly the same as in the preceding sentence. The syntactic relations between these three words, however, are different in the two sentences. The relationship between "brown" and "leaves" in the first sentence is quite different from the relationship between "Brown" and "leaves" in the second sentence. The relationship between "leaves" and "signal" is likewise different in the two sentences.

The task you are being asked to perform involves the ranking of the complexity of the syntactic relations between pairs of words such as those above. If you have any questions concerning the fact that identical groups of words can enter into different syntactic relations in different sentences, please ask them now.

On the following page you will find a list of ten sentences. In each sentence, you will find two underlined words. You will note that there are two sets of word pairs: "lost-contact" and "past (passed)-quick." Each word pair enters into different types of syntactic relations within the different sentences in which it appears. Your task is to rank the COMPLEXITY of the syntactic relations between the underlined words, beginning with the least complex and progressing to the most complex syntactic relationship. You are to rank all ten sentences in this manner. Since it is possible that each of the word pairs could enter into identical syntactic relationships, it is possible to have tie rankings. If you wish to give tie rankings to any of the sentences, assign the same numerical rank to each of the sentences entering into the tie. For instance, if you have ranked five sentences and consider the next two as being identical in complexity, assign a rank of six to each of the two sentences. The next rank would then be eight, since the seventh ranking is included in the tie ranking labeled six.

To rank these sentences in terms of the complexity of the syntactic relationship between the underlined words, please observe the following instructions. First, choose the sentence which you feel contains the least complex syntactic relation between the underlined words. Copy this sentence on the first line of the response sheet and mark this sentence number one. The sentence which contains the word pair having the next least complex relationship is then located and copied on the second line of the response sheet. This sentence is labeled number two. The remaining sentences are ranked in the same manner. Remember, tie rankings must be assigned the same number.

After you complete the ranking of the ten sentences, check your work to make sure you are satisfied with the rankings as they stand. If you wish to make any changes in the rankings, please do so on another response sheet. More response sheets are available if needed. You will be given all the time you need to complete the task, so there is no need to hurry. When you have completed the task, please remain quietly seated and do not disturb those individuals who have not yet finished.

REMEMBER, YOU ARE RANKING THESE SENTENCES IN TERMS OF THE COMPLEXITY OF THE SYNTACTIC RELATIONS BETWEEN THE UNDERLINED PAIRS OF WORDS IN EACH SENTENCE.

Are there any questions?

Sentences

A. They passed quickly through the state.

B. Even though the battle was lost, contact had been established with the rear guard.

C. He lost contact with reality.

D. The man they passed quickly turned to hide his face.

E. As in the past, quick departure is the rule.

F. If you are lost, contact the nearest policeman.

G. His present dullness contrasted sharply with his past quick thinking.

H. The lost contact lens case was finally found.

I. The voice from the past quickened his pulse.

J. The team that lost contacted the authorities.

APPENDIX D

INSTRUCTIONS AND RESPONSE FORM FOR THE SPEAKER ATTRIBUTE RANKING

INSTRUCTIONS: Sentence #____

In the following task, you will be asked to rank four speakers on the following attributes: (1) QUALITY, (2) FLUENCY, (3) NATURALNESS, (4) RATE, (5) PRECISION and (6) OVERALL EFFECTIVENESS OF SPEAKER.

1. QUALITY - A rank of one means the voice has the _most_ desirable quality, while a rank of four means the voice has the _least_ desirable quality.

   Speaker A __ Speaker B __ Speaker C __ Speaker D __

2. FLUENCY - A rank of one means the speaker is the _most_ fluent, while a rank of four means the speaker is the _least_ fluent.

   Speaker A __ Speaker B __ Speaker C __ Speaker D __

3. NATURALNESS - A rank of one means the speaker sounds the _most_ natural, while a rank of four means the speaker sounds the _least_ natural, least natural meaning artificial or "put on."

   Speaker A __ Speaker B __ Speaker C __ Speaker D __

4. RATE - A rank of one means the speaker has the _best_ overall speaking rate (not too fast and not too slow), while a rank of four means the speaker has the _worst_ overall speaking rate.

   Speaker A __ Speaker B __ Speaker C __ Speaker D __

5. PRECISION - A rank of one refers to the speaker with the _best_ precision, that is, he is neither over nor under precise. A rank of four refers to the speaker with the _worst_ precision (either over precise or imprecise).

   Speaker A __ Speaker B __ Speaker C __ Speaker D __

6. OVERALL EFFECTIVENESS OF SPEAKER - A rank of one means the speaker is the _most_ effective and a rank of four means the speaker is the _least_ effective.

   Speaker A __ Speaker B __ Speaker C __ Speaker D __

APPENDIX E

GENERATION OF THE TEST TAPES

The five stimulus sentences were recorded in an IAC series 1200 sound treated room through a Bruel and Kjaer 4134 microphone and amplifier assembly coupled to a full track Ampex 350 tape recorder. All sentences were recorded full track at a speed of 15 inches per second.

These sentences were then played back at a speed of 7-1/2 ips through a manual switch into a two-channel dubbing recorder (also operating at 7-1/2 ips) such that the output from the switch was recorded through Channel A or B of the dubbing recorder (see Figure 18). Initially, the output of the switch was connected to Channel A of the dubbing recorder. When the silent interval between "lost" and "contact" had been located auditorally (earphones) and visually (VU meter), the signal was manually switched from Channel A to B on the dubbing recorder. All five stimulus sentences were dubbed in this fashion so that the dubbed two-track tape contained the portion of the sentence up to and including the word "lost" on Track A and the remainder of the sentence on Track B.

Each "switched" two-track tape was then played back and viewed on a dual channel storage oscilloscope. The Track A and B signals were viewed individually and together, and monitored auditorally, to determine if the switch actually did take place in the silent

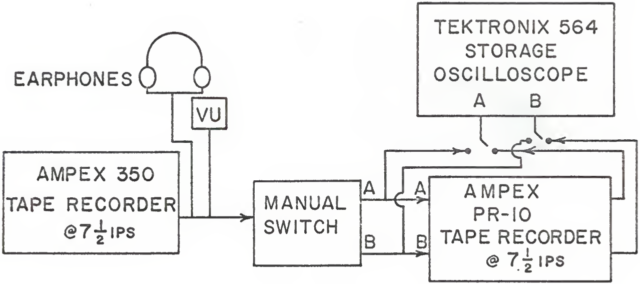

Figure 18. Block diagram of instrumentation used in the generation of the two-track test tape.

interval between the words "lost" and "contact." When the point of switching was found to occur during a portion of the speech signal, the sentence was redubbed until the switch was made within the silent interval separating the two parts of the sentence. In no case was the switching transient perceptible auditorally.

In the final stage of the recording procedure, the two-track tape containing the five stimulus sentences was played back at a real-time speed of 15 ips and redubbed at a speed of 7-1/2 ips (the speed at which the modified playback recorder used in the experimental procedure operated). This redubbing proceeded as follows.

The starting points on both the "playback" tape and the "dubbing" tape were marked. Both recorders were started simultaneously and the Track A signal recorded. Both tapes were then rewound to the starting points and the "playback" tape was then further advanced another 15-20 inches. Again both recorders were started simultaneously and the Track B signal recorded. Since the "playback" tape had been advanced 15-20 inches for this Track B dubbing, this signal was physically advanced by about one second (since the playback recorder operated at 15 ips) with respect to the original temporal relations between the Track A and B signals.

This temporal advancement of the Track B signal had the effect of allowing a subject to adjust the pause duration between "lost" and "contact" to a value less than that present in the original recording, even going to absolute zero duration or word overlap if a subject so desired. If the temporal relations between Track A and B had been left in their originally recorded state, the minimum value a subject could achieve in adjusting the duration of a silent interval would have been that value corresponding to the original duration of that interval, plus the duration represented by the physical separation of the two playback heads.

Each recorded sentence in which the Track B signal had been temporally advanced with respect to the Track A signal was made into an individual tape loop. This tape loop was then used to record approximately 150 repetitions of each stimulus sentence with a five-second interval between repetitions. The repetitions for an individual sentence were recorded on a single 1200-foot reel. Since the stimulus sentence on each reel of tape repeated itself every five seconds, the subject could be presented with as many repetitions of the sentence as needed to make the pause duration adjustment required in the experimental tasks. No more than 80 repetitions were used by any one subject in any of the experimental tasks.

Each test tape was played back through the variable delay system used in the experimental session and the pause duration between the words "lost" and "contact" adjusted until it approximated the original pause duration. Each reassembled sentence was then played into a calibrated Bruel and Kjaer graphic level recorder to obtain some relative intensity measures. The resultant graphic measures of intensity were compared to similarly obtained graphic measures of intensity from the original recordings by superimposing these two visual records. While there was a slight decrease in overall intensity from the original to the dubbed stimulus sentences, there was no difference evident between the relative intensities of the Track A and B signals from the original to the dubbed recording.

APPENDIX F

PAUSE DURATION DATA

Table 13. Pause duration data for replication one.*

	Pause Detection Task					Fluent Pause Task					Hesitation Pause Task				
	Syntactic Complexity Level					Syntactic Complexity Level					Syntactic Complexity Level				
Subject	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
1	25	87	-25	-27	-47	222	232	157	145	155	295	265	225	365	375
2	65	32	-97	-40	-45	97	147	87	232	-105	92	195	160	250	290
3	10	-95	30	-42	60	132	100	40	155	225	315	192	520	205	595
4	90	72	-87	-12	-67	142	172	117	375	232	175	190	270	477	345
5	92	97	-155	-15	-165	145	147	110	365	235	190	185	270	550	685
6	107	77	52	120	-117	100	150	75	150	40	260	250	425	580	485
7	87	-62	-147	-100	-150	137	77	135	305	200	205	217	640	1625	1375
8	12	-160	-25	-125	-62	132	197	320	257	230	130	317	392	1470	910
9	37	37	-12	-80	-147	205	67	115	295	400	200	275	520	710	2650
10	132	74	240	-22	47	167	122	340	515	215	145	357	680	710	280
11	135	-10	-50	70	130	165	162	165	460	350	425	370	435	1060	1120
12	95	77	60	145	-32	105	65	60	190	2	207	197	190	430	535

*Durations in msec.

Table 14. Pause duration data for replication two.*

	Pause Detection Task					Fluent Pause Task					Hesitation Pause Task				
	Syntactic Complexity Level					Syntactic Complexity Level					Syntactic Complexity Level				
Subject	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
1	60	117	-2	2	72	272	210	165	187	155	260	260	260	305	400
2	2	-22	-170	-32	-52	180	80	75	175	47	275	280	142	590	270
3	120	92	-10	135	100	82	35	87	127	190	210	102	315	420	215
4	72	80	22	-22	-107	192	187	187	300	455	180	155	285	390	510
5	47	80	-67	-110	-2	142	92	132	275	195	227	235	270	297	355
6	77	75	-47	25	-110	135	150	117	122	97	205	245	635	435	510
7	40	-45	-32	92	-135	117	155	137	232	310	275	225	800	1600	1775
8	137	78	-35	92	72	162	202	155	232	195	290	517	500	1580	415
9	30	-32	-205	-197	-160	150	182	275	320	390	217	262	560	1000	2750
10	65	55	-5	-82	-132	140	80	495	430	225	372	275	585	650	440
11	177	-102	-65	-7	-60	200	147	215	355	245	335	355	780	1075	1110
12	90	27	-52	97	-87	90	60	107	205	157	292	177	290	380	615

---

*Durations in msec.

Table 15. Pause duration data for replication three.*

	Pause Detection Task					Fluent Pause Task					Hesitation Pause Task				
	Syntactic Complexity Level					Syntactic Complexity Level					Syntactic Complexity Level				
Subject	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
1	-10	137	92	-158	55	182	175	92	172	90	347	445	427	515	500
2	22	-20	32	50	5	165	62	70	230	60	315	315	67	680	440
3	17	-60	-7	145	-110	110	70	60	115	60	205	205	285	107	580
4	130	57	-55	-20	-7	160	210	190	227	200	245	152	255	270	730
5	122	132	27	-27	-45	155	182	127	322	250	352	270	372	620	425
6	87	57	-5	142	-42	120	140	122	172	135	232	315	525	627	575
7	60	67	5	50	-62	250	155	117	235	290	195	225	535	1440	1350
8	47	80	167	-75	-82	130	185	157	630	197	527	442	460	1580	1300
9	17	-22	-182	-85	20	57	147	235	337	440	280	200	620	930	2200
10	127	90	-7	37	-40	110	215	582	290	267	200	395	840	590	245
11	100	0	-65	-92	-17	210	135	225	450	375	530	435	825	1380	1045
12	100	20	-25	0	-50	122	37	190	155	220	212	217	525	525	1010

*Durations in msec.

APPENDIX G

MEAN PAUSE DURATION AND STANDARD DEVIATION AT EACH LEVEL OF SYNTACTIC COMPLEXITY

Table 16. Mean and standard deviation of pause detection thresholds.*

	Level of Syntactic Complexity									
	1		2		3		4		5	
Subject	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
1	25	29	114	21	22	51	-61	70	27	53
2	30	26	-4	25	-78	83	-7	41	-31	25
3	49	50	-21	81	4	18	79	86	17	91
4	97	24	70	10	-40	46	-18	4	-60	41
5	87	31	3	22	-66	74	-51	42	-71	68
6	90	12	70	9	0	41	96	51	-90	34
7	62	19	-13	57	-58	65	14	82	-116	38
8	65	53	-1	112	36	93	-36	92	-24	68
9	28	8	-6	30	-133	86	-121	54	-96	82
10	8	30	73	14	76	115	-22	49	-42	73
11	37	31	-37	46	-60	7	-10	66	18	81
12	95	4	41	25	-6	48	81	60	-56	23

*Durations in msec.

Table 17. Mean and standard deviation of fluent pause durations.*

	Level of Syntactic Complexity									
	1		2		3		4		5	
Subject	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
1	225	37	206	23	138	33	168	17	133	31
2	147	36	96	37	77	7	212	26	101	75
3	108	20	68	27	62	19	132	17	158	71
4	165	20	190	15	165	33	301	60	296	113
5	147	6	140	37	123	9	321	37	227	23
6	118	14	147	4	105	21	148	20	91	39
7	168	59	129	37	130	9	257	34	267	48
8	141	15	196	8	211	77	373	182	207	16
9	137	61	132	48	208	68	317	17	410	22
10	139	23	139	56	472	100	412	93	236	23
11	192	19	148	11	202	26	422	47	323	56
12	106	13	54	12	119	54	183	21	126	91

*Durations in msec.

Table 18. Mean and standard deviation of hesitation pause durations.*

Subject	Level of Syntactic Complexity 1 Mean	1 S.D.	2 Mean	2 S.D.	3 Mean	3 S.D.	4 Mean	4 S.D.	5 Mean	5 S.D.
1	301	36	323	86	304	88	395	88	425	54
2	227	97	263	50	123	40	507	185	333	76
3	243	51	166	46	373	104	244	130	463	175
4	200	32	266	17	270	12	379	85	528	157
5	256	59	230	35	304	48	489	138	488	142
6	232	22	270	32	528	86	547	82	523	38
7	225	36	222	4	658	109	1555	82	1500	195
8	316	163	422	83	451	45	1543	52	875	362
9	232	34	246	33	567	41	880	124	2533	239
10	239	97	342	50	702	105	650	49	322	85
11	430	80	387	35	683	177	1172	147	1092	33
12	237	39	197	16	335	140	445	60	720	207

*Durations in msec.

APPENDIX H

MEAN PAUSE DURATION AND STANDARD DEVIATION FOR EACH SYNTACTIC COMPLEXITY LEVEL AND TASK

Table 19. Mean and standard deviation of pause detection threshold, fluent pause and hesitation pause durations at each level of syntactic complexity.*

Task	Syntactic Complexity Level	Replication 1		Replication 2		Replication 3	
		Mean	S.D.	Mean	S.D.	Mean	S.D.
Pause Detection Threshold	1	74	42	76	46	68	47
	2	19	79	33	65	44	59
	3	-18	103	-56	64	-2	81
	4	-11	79	-1	93	-3	89
	5	-50	87	-51	86	-32	43
Fluent Pause	1	145	37	155	49	147	49
	2	136	50	131	57	142	56
	3	143	90	178	109	180	132
	4	287	117	246	89	277	138
	5	181	136	221	111	215	114
Hesitation Pause	1	219	87	261	54	303	114
	2	250	64	257	100	301	101
	3	394	166	451	210	478	215
	4	702	437	726	455	772	447
	5	804	646	780	724	867	527

*Durations in msec.

BIBLIOGRAPHY

Aaronson, D. and Markowitz, N. (1967). "Recall of Compressed Speech," Bulletin of the Center for Rate Controlled Recordings, Vol. 1, No. 2.

Agnello, J. (1963). "A Study of Intra- and Inter-Phrasal Pauses and Their Relationship to the Rate of Speech," Ohio State Univ. Dissertation, 131 p.

Black, J., Tosi, O., Singh, S. and Takefuta, Y. (1966). "A Study of Pauses in Oral Reading of One's Native Language and English," Language and Speech, Vol. 9, pp. 17-25.

Bloomfield, L. (1933). Language, New York: Holt, Rinehart and Winston.

Boomer, D. and Dittman, A. (1962). "Hesitation Pauses and Juncture Pauses in Speech," Language and Speech, Vol. 5, pp. 215-226.

Carrell, J. and Tiffany, W. (1960). Phonetics: Theory and Application to Speech Improvement, New York: McGraw Hill, Inc.

Cooper, F. S., Delattre, P., Liberman, A., Borst, J. and Gerstman, L. (1952). "Some Experiments on the Perception of Synthetic Speech Sounds," J. Acoust. Soc. Am., Vol. 24, pp. 597-606.

Goldman-Eisler, F. (1961a). "A Comparative Study of Two Hesitation Phenomena," Language and Speech, Vol. 4, pp. 18-26.

Goldman-Eisler, F. (1961b). "The Distribution of Pause Durations in Speech," Language and Speech, Vol. 4, pp. 232-237.

Goldman-Eisler, F. (1968). Psycholinguistics: Experiments in Spontaneous Speech, New York: Academic Press.

Hargreaves, W. A. and Starkweather, J. (1959). "Collection of Temporal Data with the Duration Tabulator," J. of Ex. Analysis of Behav., Vol. 2, pp. 179-183.

Jones, D. (1962). An Outline of English Phonetics, Cambridge: W. Heffer & Sons.

Judson, L. and Weaver, A. (1965). Voice Science, New York: Appleton-Century-Crofts.

Lieberman, P. (1967). Intonation, Perception and Language, Cambridge: The MIT Press.

Lindquist, E. F. (1956). Design and Analysis of Experiments in Psychology and Education, Boston: Houghton Mifflin Company,

Lounsbury, F. (1954). "Sequential Psycholinguistics," in Psycholinguistics, A Survey of Theory and Research Problems (C. Osgood and T. Sebeok, eds.), Baltimore: Waverly Press, Inc., pp. 93-125.

Lynch, G. (1934). "A Phonophotographic Study of Trained and Untrained Voices Reading Factual and Dramatic Material," Arch. of Speech, Vol. 1, pp. 9-25.

Maclay, H. and Osgood, C. (1959). "Hesitation Phenomena in Spontaneous English Speech," Word, Vol. 15, pp. 19-44.

Mattingly, I. (1966). "Synthesis by Rule of Prosodic Features," Language and Speech, Vol. 9, pp. 1-13.

Miller, G. and Chomsky, N. (1963). "Finitary Models of Language Users," in Handbook of Mathematical Psychology (R. Luce, R. Bush and E. Galanter, eds.), Vol. II, pp. 419-492.

Pike, K. L. (1945). The Intonation of American English, Ann Arbor: U. of Michigan Press.

Scholes, Robert J. (1968). "On the Spoken Disambiguation of Superficially Ambiguous Sentences," Bell Telephone Laboratories Technical Memorandum, July.

Small, A. and Campbell, R. (1962). "Temporal Differential Sensitivity for Auditory Stimuli," Am. J. of Psych., Vol. 75, No. 3, pp. 401-410.

Steiner, S. (1964). "Apparent Duration of Auditory Stimuli," State University of Iowa Dissertation, 162 p.

Trager, G. and Smith, H. (1951). An Outline of English Structure, SIL: Occassional Papers, 3: Norman, Oklahoma.

Winer, B. J. (1962). Statistical Principles in Experimental Design, New York: McGraw-Hill Book Co.

BIOGRAPHICAL SKETCH

Kenneth F. Ruder was born November 27, 1939, at Hays, Kansas. In June, 1957, he was graduated from St. Joseph's Military Academy. In June, 1962, he received the degree of Bachelor of Arts with a major in Speech Pathology from Fort Hays Kansas State Teachers College. In September, 1962, he enrolled in the University of Wisconsin Graduate School. He held a Traineeship from the Vocational Rehabilitation Administration until August, 1963, when he received the degree of Master of Science with a major in Speech Pathology and Audiology. From 1964 until 1966, he served as Director of Speech Pathology Services at Harmarville Rehabilitation Center, Pittsburgh, Pennsylvania. In 1966, he enrolled in the Graduate School of the University of Florida, and from that time until the present, he has pursued his work toward the degree of Doctor of Philosophy.

Kenneth F. Ruder is married to the former Charlotte C. Collins and is the father of two daughters. He is a member of Sigma Xi and the Acoustical Society of America.

This dissertation was prepared under the direction of the chairman of the candidate's supervisory committee and has been approved by all members of that committee. It was submitted to the Dean of the College of Arts and Sciences and to the Graduate Council, and was approved as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

August, 1969

E. Ruffin Jones
Dean, College of Arts and Sciences

Dean, Graduate School

Supervisory Committee:

Paul Jensen
Chairman

Harry Hollis PhD

Robert J. Scholes

John F. Brandt

John Algeo